

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2002, 15:42:29 ; Search time 12.89 Seconds  
(without alignments)  
115.590 Million cell updates/sec

Title: US-09-829-481-4  
Perfect score: 339  
Sequence: 1 MKSIALIFIVLVAFCILEDGC.....IRRRGGFCRGTFRTTCVCYR 61

Scoring table: BioSUM62  
Gapext 0.5

Searched: 231628 seqs., 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0\*

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:  
 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:  
 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:  
 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:  
 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:  
 5: /cgn2\_6/ptodata/2/iaa/8CUTUS\_COMB.pep:  
 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:  
 \* Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	68.5	20.2	72	4	US-09-003-198A-17		Sequence 17, Appl
2	66	19.5	40	1	US-08-385-375-16		Sequence 16, Appl
3	66	19.5	40	1	US-08-385-375-19		Sequence 39, Appl
4	64.5	19.0	80	4	US-09-489-20		Sequence 20, Appl
5	64	18.9	74	4	US-09-442-631-4		Sequence 4, Appl
6	63.5	18.7	768	1	US-08-454-455-6		Sequence 4, Appl
7	63.5	18.7	769	1	US-08-454-455-6		Sequence 6, Appl
8	61.5	18.1	80	1	US-08-377-687-49		Sequence 49, Appl
9	61.5	18.1	80	2	US-08-777-192-49		Sequence 49, Appl
10	61.5	18.1	80	4	US-08-971-982-49		Sequence 49, Appl
11	60.5	17.8	81	4	US-09-053-021-4		Sequence 4, Appl
12	60.5	17.8	81	4	US-09-053-021-4		Sequence 9, Appl
13	60.5	17.8	652	2	US-08-751-305-2		Sequence 2, Appl
14	60	17.7	79	1	US-08-627-706-15		Sequence 15, Appl
15	60	17.7	79	4	US-09-103-489-15		Sequence 15, Appl
16	59.5	17.6	42	2	US-08-751-305-3		Sequence 3, Appl
17	59.5	17.6	80	1	US-08-377-687-59		Sequence 59, Appl
18	59.5	17.6	80	1	US-08-777-192-59		Sequence 59, Appl
19	59.5	17.6	80	4	US-08-971-982-59		Sequence 59, Appl
20	59	17.4	204	1	US-08-652-859-2		Sequence 2, Appl
21	59	17.4	204	2	US-08-919-106-2		Sequence 2, Appl
22	59	17.4	204	2	US-09-153-751-2		Sequence 2, Appl
23	58.5	17.3	75	1	US-08-289-458-2		Sequence 2, Appl
24	58.5	17.3	75	2	US-08-761-549-2		Sequence 2, Appl
25	58.5	17.3	75	4	US-09-127-646-2		Sequence 2, Appl
26	57.5	17.0	221	2	US-08-925-708-1		Sequence 1, Appl
27	57	16.8	86	1	US-08-149-839B-9		Sequence 14, Appl

## ALIGNMENTS

RESULT 1  
US-09-003-198A-17  
; Sequence 17, Application US/09003198A  
; Patent No. 6316407  
; GENERAL INFORMATION:  
 ; APPLICANT: Liang, Jihong Maganalal  
 ; APPLICANT: Shah, Dilip Maganalal  
 ; APPLICANT: Wu, Yonnie S.  
 ; APPLICANT: Rosenberger, Cindy A.  
 ; APPLICANT: Hakimi, Salim  
 ; TITLE OF INVENTION: Antifungal Polypeptide and Methods for Controlling Plant Pathogenic Fungi  
 ; NUMBER OF SEQUENCES: 19  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Arnold White & Durkee  
 ; STREET: P.O. Box 4433  
 ; CITY: Houston  
 ; STATE: Texas  
 ; COUNTRY: USA  
 ; ZIP: 77210  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/003,198A  
 ; FILING DATE: 07-JAN-1998  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Patterson, Melinda L.  
 ; REGISTRATION NUMBER: 33,062  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (713) 787-1400  
 ; TELEFAX: (713) 787-1440  
 ; INFORMATION FOR SEQ ID NO: 17:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 72 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: linear  
 ; US-09-003-198A-17  
 Query Match Score 20.28;  
 Best Local Similarity 28.4%; Pred. No. 0/26;  
 Matches 19; Conservative 11; Mismatches 24; Indels 13; Gaps 3;  
 QY 2 K\$IAIIIVLVAFCLIEDGIVEAGEGCPFNACK-----CHRHKSKIRRGG--GFCCR 50

Db 4 KSLACLSFLLVLVLFVAQEIVVSEANTCENLAGSYKGVCFGGCDRHCR - QEGATSGRCR 61  
 Qy 51 GTFRITC 57  
 Db 62 DDFRCWC 68

RESULT 2  
 US-08-385-375-16  
 Sequence 16, Application US/08385375  
 Patent No. 5631144  
 GENERAL INFORMATION:  
 APPLICANT: LEMOINE, Yves  
 APPLICANT: NGUYEN, Martine  
 APPLICANT: ACHSTETTER, Tilman  
 APPLICANT: REICHHART, Jean-Marc  
 TITLE OF INVENTION: APPLICATION OF NOVEL DNA FRAGMENTS AS A  
 TITLE OF INVENTION: APPLICATION OF NOVEL PEPTIDE FOR THE SECRETION OF  
 TITLE OF INVENTION: SEQUENCE CODING FOR A SIGNAL PEPTIDE FOR THE SECRETION OF  
 TITLE OF INVENTION: MATURE PROTEINS BY RECOMBINANT YEASTS, EXPRESSION  
 TITLE OF INVENTION: CASSETTES, TRANSFORMED YEASTS AND CORRESPONDING PROCESS FOR  
 TITLE OF INVENTION: THE PREPARATION OF PROTEINS  
 NUMBER OF SEQUENCES: 40  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Burns, Doane, Swecker & Mathis  
 STREET: P.O. Box 1404  
 CITY: Alexandria  
 STATE: Virginia  
 COUNTRY: United States  
 ZIP: 22313-1404  
 COMPUTER READABLE FORM:  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/385,375  
 FILING DATE: 2002-08-08  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/178,356  
 FILING DATE: 04-JAN-1994  
 APPLICATION NUMBER: FR 89/05687  
 FILING DATE: 28-APR-1994  
 PRIOR APPLICATION NUMBER:  
 APPLICATION NUMBER: WO PCT/FR90/00306  
 FILING DATE: 27-APR-1990  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/178,356  
 FILING DATE: 04-JAN-1994  
 APPLICATION NUMBER: FR 89/05687  
 FILING DATE: 28-APR-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: WO PCT/FR90/00306  
 FILING DATE: 27-APR-1990  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Crane-Feury, Sharon E  
 FILING DATE: 04-JAN-1994  
 APPLICATION NUMBER: 36,113  
 FILING DATE: 28-APR-1994  
 REFERENCE/DOCKET NUMBER: 017753-009  
 TELEPHONE: (703) 836-6620  
 REFERENCE/DOCKET NUMBER: 017753-009  
 TELEPHONE: (703) 836-6620  
 REFERENCE/DOCKET NUMBER: 017753-009  
 TELEPHONE: (703) 836-2021  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 40 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-385-375-16

Query Match 19.5%; Score 66; DB 1; Length 40;  
 Best Local Similarity 39.4%; Pred. No. 0.28; DB 1; Length 40;  
 Matches 13; Conservative 2; Mismatches 16; Indels 2; Gaps 1;

Qy 27 GCPFNAGKCHRHKCKSIRRRGGFCRGTFRTTCYC 59  
 Db 8 GTGINHSACAACHLRLRGNGGCGCNG - KGVCYC 38

RESULT 3  
 Sequence 39, Application US/08385375  
 Patent No. 5631144  
 GENERAL INFORMATION:  
 APPLICANT: LEMOINE, Yves  
 APPLICANT: NGUYEN, Martine  
 APPLICANT: ACHSTETTER, Tilman  
 APPLICANT: REICHHART, Jean-Marc  
 TITLE OF INVENTION: APPLICATION OF NOVEL DNA FRAGMENTS AS A  
 TITLE OF INVENTION: APPLICATION OF NOVEL PEPTIDE FOR THE SECRETION OF  
 TITLE OF INVENTION: SEQUENCE CODING FOR A SIGNAL PEPTIDE FOR THE SECRETION OF  
 TITLE OF INVENTION: MATURE PROTEINS BY RECOMBINANT YEASTS, EXPRESSION  
 TITLE OF INVENTION: CASSETTES, TRANSFORMED YEASTS AND CORRESPONDING PROCESS FOR  
 TITLE OF INVENTION: THE PREPARATION OF PROTEINS  
 NUMBER OF SEQUENCES: 40  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Burns, Doane, Swecker & Mathis  
 STREET: P.O. Box 1404  
 CITY: Alexandria  
 STATE: Virginia  
 COUNTRY: United states  
 ZIP: 22313-1404  
 COMPUTER READABLE FORM:  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/385,375  
 FILING DATE: 2002-08-08  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/178,356  
 FILING DATE: 04-JAN-1994  
 APPLICATION NUMBER: FR 89/05687  
 FILING DATE: 28-APR-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: WO PCT/FR90/00306  
 FILING DATE: 27-APR-1990  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Crane-Feury, Sharon E  
 FILING DATE: 04-JAN-1994  
 APPLICATION NUMBER: 36,113  
 FILING DATE: 28-APR-1994  
 REFERENCE/DOCKET NUMBER: 017753-009  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703) 836-6620  
 TELEFAX: (703) 836-2021  
 INFORMATION FOR SEQ ID NO: 39:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 40 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-385-375-39

Query Match 19.5%; Score 66; DB 1; Length 40;  
 Best Local Similarity 39.4%; Pred. No. 0.28; DB 1; Length 40;  
 Matches 13; Conservative 2; Mismatches 16; Indels 2; Gaps 1;

Qy 27 GCPFNAGKCHRHKCKSIRRRGGFCRGTFRTTCYC 59  
 Db 8 GTGINHSACAACHLRLRGNGGCGCNG - KGVCYC 38

RESULT 4  
 Sequence 20, Application US/09103489  
 Patent No. 6215048  
 GENERAL INFORMATION:  
 APPLICANT: Liang, Jihong  
 APPLICANT: Shah, Dilip M.  
 APPLICANT: Wu, Yonnie S.

APPLICANT: Rosenberger, Cindy A.  
 TITLE OF INVENTION: Antifungal Polypeptide and Methods for Controlling Plant Pathogenic Fungi  
 NUMBER OF SEQUENCES: 200  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Charles E. Cohen, Monsanto Company, BB4F STREET: 700 Chesterfield Village Parkway No. 6215048th STATE: St. Louis CITY: Missouri COUNTRY: USA ZIP: 63198 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: PatentIn Release #1.0, Version #1.30 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/103, 489 FILING DATE: 24-JUN-1998 CLASIFICATION: 800 ATTORNEY/AGENT INFORMATION:  
 NAME: Cohen, Charles E.  
 REGISTRATION NUMBER: 34, 565 REFERENCE/DOCKET NUMBER: 38-21 (10700)A TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (314) 537-6224 TELEFAX: (314) 537-6047 INFORMATION FOR SEQ ID NO: 20:  
 LENGTH: 80 amino acids SEQUENCE CHARACTERISTICS:  
 TYPE: amino acid STRANDEDNESS:  
 TOPOLOGY: linear MOLECULE TYPE: peptide

US-09-103-489-20

Query Match 18.9%; Score 64; DB 4; Length 74;  
 Best Local Similarity 35.8%; Pred. No. 0.88; Mismatches 23; Indels 12; Gaps 5;

QY 2 KSIAILIVLVAFC--TLEDGIVEAGFGCPFN--AGKC--HRHCKSIRRGGFCRGRT--- 52  
 Db 6 KVATIFLMMKVATDMAEAKICAEALSG --NFKGLCLSSRDCGNVRREGFTDSCIG 62

QY 53 FRTTCVC 59  
 Db 63 FRLQCFC 69

RESULT 6  
 US-08-454-455-4  
 Sequence 4, Application US/08454455  
 Patent No. 5633601  
 GENERAL INFORMATION:  
 APPLICANT: Moyle, Matthew  
 APPLICANT: McLean, John W.  
 TITLE OF INVENTION: NOVEL BETA INTEGRIN SUBUNIT  
 NUMBER OF SEQUENCES: 9  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Genentech, Inc.  
 STREET: 460 Point San Bruno Blvd  
 CITY: South San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94080  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5 inch, 720 Kb floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: WinPatin (Genentech)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/454,455  
 FILING DATE: 30-MAY-1995  
 CLASSIFICATION: 530  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 08/193989  
 FILING DATE: 09-FEB-1994  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 08/004142  
 FILING DATE: 13-JAN-1993  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 07/670607  
 FILING DATE: 14-MAR-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Lee, Wendy M.  
 REGISTRATION NUMBER: 00,000  
 REFERENCE/DOCKET NUMBER: P0699C2D2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415/225-1994  
 TELEFAX: 415/952-9881  
 TELEX: 910/371-7168  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 768 amino acids  
 TYPE: Amino Acid  
 TOPOLOGY: Linear  
 US-08-454-455-4

Query Match 18.7%; Score 63.5; DB 1; Length 768;  
 Best Local Similarity 24.7%; Pred. No. 10; Mismatches 8; Indels 39; Gaps 5;

QY 14 FCILEDGIVLVAEGFCPN-----AGKCH----- 51  
 Db 545 YCEKDD----FSCPIFHGSILCAGHEACRGCQCFSGWEGDRCCOPSSAAQHC -VNS 596

QY 45 RGGFRCRTFRITCVCYR 61

US-09-442-631-4

APPLICANT: SHIN, BYONCHUL  
 APPLICANT: OH, BOUNG-JUN  
 APPLICANT: KO, MOON KYUNG  
 APPLICANT: CHUNG, CHANG HO  
 TITLE OF INVENTION: SMALL AND CYSTEINE RICH ANTIFUNGAL DEFENSIN AND THIONIN-LIKE PROTEIN GENES HIGHLY EXPRESSED IN THE FILE REFERENCE: 1942/44  
 CURRENT APPLICATION NUMBER: US/09/442,631  
 CURRENT FILING DATE: 1999-11-18  
 NUMBER OF SEQ ID NOS: 4  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 4  
 LENGTH: 74  
 TYPE: PRT  
 ORGANISM: Capsicum annuum  
 US-09-442-631-4

APPLICANT: CAMMUE, BRUNO P. A.  
 APPLICANT: OSBORN, RUPERT W.  
 APPLICANT: REES, SARAH B.  
 APPLICANT: TERRAS, FRANKY R. G.  
 APPLICANT: VANDERLEYDEN, JOZEF F.

TITLE OF INVENTION: BIOCIDAL PROTEINS  
 NUMBER OF SEQUENCES: 59

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: CUSHMAN DARBY & CUSHMAN  
 STREET: 1100 NEW YORK AVENUE, N.W.  
 CITY: WASHINGTON D.C.

STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20005

COMPUTER READABLE FORM:  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/377,687  
 FILING DATE: 09-04-2005  
 CLASSIFICATION: 800  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US/08/454,455  
 FILING DATE: 30-MAY-1995  
 CLASSIFICATION: 530  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 08/193989  
 FILING DATE: 09-FEB-1994  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 08/004142  
 FILING DATE: 13-MAR-1993  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 07/670607  
 FILING DATE: 14-MAR-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Lee, Wendy M.  
 REGISTRATION NUMBER: 00,000  
 REFERENCE/DOCKET NUMBER: P0699CDD2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415/225-1994  
 TELEFAX: 415/952-9881  
 TELEX: 910/731-7168  
 INFORMATION FOR SEQ ID NO: 6:  
 LENGTH: 769 amino acids  
 SEQUENCE CHARACTERISTICS:  
 TYPE: Amino Acid  
 TOPOLOGY: Linear  
 US-08-454-455-6

Query Match 7  
 Sequence 6, Application US/08454455  
 Patent No. 5635601  
 GENERAL INFORMATION:  
 APPLICANT: McLean, John W.  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Genentech, Inc.  
 STREET: 460 Point San Bruno Blvd  
 CITY: South San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94080  
 COMPUTER READABLE FORM:  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: WinPatin (Genentech)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/454,455  
 FILING DATE: 30-May-1995  
 CLASSIFICATION: 530  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 08/193989  
 FILING DATE: 09-FEB-1994  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 08/004142  
 FILING DATE: 13-MAR-1993  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 07/670607  
 FILING DATE: 14-MAR-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Lee, Wendy M.  
 REGISTRATION NUMBER: 00,000  
 REFERENCE/DOCKET NUMBER: P0699CDD2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415/225-1994  
 TELEFAX: 415/952-9881  
 TELEX: 910/731-7168  
 INFORMATION FOR SEQ ID NO: 6:  
 LENGTH: 769 amino acids  
 SEQUENCE CHARACTERISTICS:  
 TYPE: Amino Acid  
 TOPOLOGY: Linear  
 US-08-454-455-6

Query Match 8  
 Sequence 49, Application US/08377687  
 Patent No. 553855  
 GENERAL INFORMATION:  
 APPLICANT: BROEKART, WILLEM F.  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: CUSHMAN DARBY & CUSHMAN  
 STREET: 1100 NEW YORK AVENUE, N.W.  
 CITY: WASHINGTON D.C.

RESULT 7  
 Sequence 6, Application US/08454455  
 Patent No. 5635601  
 GENERAL INFORMATION:  
 APPLICANT: McLean, John W.  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Genentech, Inc.  
 STREET: 460 Point San Bruno Blvd  
 CITY: South San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94080  
 COMPUTER READABLE FORM:  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/377,687  
 FILING DATE: 09-04-2005  
 CLASSIFICATION: 800  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US/08/454,455  
 FILING DATE: 30-MAY-1995  
 CLASSIFICATION: 530  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 08/193989  
 FILING DATE: 09-FEB-1994  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 08/004142  
 FILING DATE: 13-MAR-1993  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 07/670607  
 FILING DATE: 14-MAR-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Lee, Wendy M.  
 REGISTRATION NUMBER: 00,000  
 REFERENCE/DOCKET NUMBER: P0699CDD2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415/225-1994  
 TELEFAX: 415/952-9881  
 TELEX: 910/731-7168  
 INFORMATION FOR SEQ ID NO: 6:  
 LENGTH: 769 amino acids  
 SEQUENCE CHARACTERISTICS:  
 TYPE: Amino Acid  
 TOPOLOGY: Linear  
 US-08-454-455-6

Query Match 9  
 Sequence 49, Application US/08777192  
 Patent No. 5824869  
 GENERAL INFORMATION:  
 APPLICANT: BROEKART, WILLEM F.  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: CUSHMAN DARBY & CUSHMAN  
 STREET: 1100 NEW YORK AVENUE, N.W.  
 CITY: WASHINGTON D.C.

RESULT 8  
 Sequence 49, Application US/08377687  
 Patent No. 553855  
 GENERAL INFORMATION:  
 APPLICANT: BROEKART, WILLEM F.  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: CUSHMAN DARBY & CUSHMAN  
 STREET: 1100 NEW YORK AVENUE, N.W.  
 CITY: WASHINGTON D.C.

COUNTRY: USA  
 ZIP: 20005  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/777,192  
 FILING DATE: 04-JAN-1993  
 CLASSIFICATION:  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US 08/002,480  
 FILING DATE: 04-JAN-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: KOKULIS, PAUL N.  
 REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-861-3000  
 TELEFAX: 202-822-0944  
 INFORMATION FOR SEQ ID NO: 49:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 80 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-777-192-49

Query Match 18.1%; Score 61.5; DB 2; Length 80;  
 Best Local Similarity 30.0%; Pred. No. 1.9;  
 Matches 21; Conservative 8; Missmatches 28; Indels 13; Gaps 4;  
 RESULT 11  
 US-09-053-021-4

Qy 4 IAIIFIVLVAFICLE-DGIVEAAGFGCPFNAG-----KCHRHKSTRR-RGGFCRG 51  
 ||::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 8 IALLFAALVLFAAFEAPTMVEAQKLCEPRTWSGVCGNNACKNQCNLEKARHGSCNY 67  
 - Qy 52 TFRT-TCYCY 60  
 Db 68 VFPAAHKCICY 77

Query Match 18.1%; Score 61.5; DB 2; Length 80;  
 Best Local Similarity 30.0%; Pred. No. 1.9;  
 Matches 21; Conservative 8; Missmatches 28; Indels 13; Gaps 4;  
 RESULT 10  
 US-08-971-982-49

Sequence 49 Application US/08971982  
 Pat. No. 6187304  
 GENERAL INFORMATION:  
 APPLICANT: BROekaert, Willem F.,  
 CAMMUE, Bruno P.A.  
 OSBORN, Rupert W.  
 REES, Sarah B.  
 TERRAS, Franky R.G.  
 VANDERLEIJDEN, JOZEF  
 TITLE OF INVENTION: BIOTICIDAL PROTEINS  
 NUMBER OF SEQUENCES: 59  
 CORRESPONDENCE ADDRESS:  
 ADDRESSSEE: CUSHMAN DARBY & CUSHMAN  
 STREET: 1100 NEW YORK AVENUE, N.W.  
 CITY: WASHINGTON  
 STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20005

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/971,982  
 FILING DATE: 17-No. 6187304-1997  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/002,480  
 FILING DATE: 04-JAN-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: KOKULIS, PAUL N.  
 REFERENCE/DOCKET NUMBER: 16-773  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-861-3000  
 TELEFAX: 202-822-0944  
 INFORMATION FOR SEQ ID NO: 49:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 80 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-971-982-49

Query Match 18.1%; Score 61.5; DB 4; Length 80;  
 Best Local Similarity 30.0%; Pred. No. 1.9;  
 Matches 21; Conservative 8; Missmatches 28; Indels 13; Gaps 4;  
 RESULT 11  
 US-09-053-021-4

Qy 4 IAIIFIVLVAFICLE-DGIVEAAGFGCPFNAG-----KCHRHKSTRR-RGGFCRG 51  
 ||::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 Qy 4 IAIIFIVLVAFICLE-DGIVEAAGFGCPFNAG-----KCHRHKSTRR-RGGFCRG 51  
 ||::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 8 IALLFAALVLFAAFEAPTMVEAQKLCEPRTWSGVCGNNACKNQCNLEKARHGSCNY 67  
 - Qy 52 TFRT-TCYCY 60  
 Db 68 VFPAAHKCICY 77

Query Match 18.1%; Score 61.5; DB 2; Length 80;  
 Best Local Similarity 30.0%; Pred. No. 1.9;  
 Matches 21; Conservative 8; Missmatches 28; Indels 13; Gaps 4;  
 RESULT 11  
 US-09-053-021-4

Sequence 4, Application US/09053021  
 Patent No. 6220785  
 GENERAL INFORMATION:  
 APPLICANT: SELSKO, Barbara  
 APPLICANT: GARCIA-RODRIGUEZ, Consuelo  
 APPLICANT: ZAMUDIO-ZUNIGA, Fernando  
 APPLICANT: BECERRIL-LUJAN, Baltazar  
 APPLICANT: POSSANI-POSTAY, Lourival D.  
 TITLE OF INVENTION: Primary Sequence and cDNA of  
 Patent No. 6220785  
 TITLE OF INVENTION: Insecticidally Effective Toxins from Scorpions of the  
 NUMBER OF SEQUENCES: 9  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Finnean, Henderson, Farabow, Garrett &  
 Dunner, L.L.P.  
 STREET: 1300 I Street, N.W.  
 CITY: Washington  
 STATE: DC  
 COUNTRY: USA  
 ZIP: 20005  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/053,021  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/017,007  
 FILING DATE: 30-APR-1996  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/848,261  
 FILING DATE: 29-APR-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Garrett, Arthur S.  
 REGISTRATION NUMBER: 20,338

REFERENCE/DOCKET NUMBER: 06899.0001-01000  
 TELECOMMUNICATION: (202)408-4000  
 TELEFAX: (202)408-4400  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
     LENGTH: 81 amino acids  
     TYPE: amino acid  
     TOPOLOGY: linear  
     MOLECULE TYPE: protein  
 US-09-053-021-4

Query Match 17.8%; Score 60.5%; DB 4; Length 81;  
 Best Local Similarity 23.4%; Pred. No. 2.4;  
 Matches 15; Conservative 10; Mismatches 28; Indels 11; Gaps 3;

QY 4 IAIIFIVLVAFCLIEDGIVEAGFCGPINA-----GKCHRCKSIRRG-GFCRGTRFT 55  
 Db 1 ITACLVLIGTVCAKEGYLVNKSTGCKYNCNLILGENKNCDMECKAKNQGSYGC--YKL 57

QY 56 TCVC 59  
 Db 58 ACWC 61

RESULT 13  
 US-08-751-305-2  
 ; Sequence 2, Application US/08751305  
 ; Patent No. 5965439

GENERAL INFORMATION:  
 ; APPLICANT: Tanner et al., Andrea J.  
 ; TITLE OF INVENTION: Host Defense Enhancement  
 ; NUMBER OF SEQUENCES: 33  
 ; CORRESPONDENCE ADDRESS:  
     ; ADDRESSEE: Fish & Richardson P.C.  
     ; STREET: 4225 Executive Square, Suite 1400  
     ; CITY: La Jolla  
     ; STATE: CA  
     ; COUNTRY: USA  
     ; 2TIP: 92037

COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent Release #1.0, version #1.30  
 ; CURRENT APPLICATION DATA:  
     ; APPLICATION NUMBER: US/08751305  
     ; FILING DATE: 18-NOV-1996  
     ; CLASSIFICATION: 514  
 ATTORNEY/AGENT INFORMATION:  
     ; NAME: Wetherell, Jr., John R.  
     ; REGISTRATION NUMBER: 31,778  
     ; REFERENCE/DOCKET NUMBER: 07305/012001  
 TELECOMMUNICATION INFORMATION:  
     ; TELEPHONE: 619/678-5070  
     ; TELEFAX: 619/678-5059  
     ; INFORMATION FOR SEQ ID NO: 2:  
         ; SEQUENCE CHARACTERISTICS:  
             ; LENGTH: 652 amino acids  
             ; TYPE: amino acid  
             ; TOPOLOGY: linear  
             ; MOLECULE TYPE: protein  
 US-08-751-305-2

RESULT 14  
 US-08-627-706-15  
 ; Sequence 15, Application US/08627706  
 ; Patent No. 5733696

GENERAL INFORMATION:  
 ; APPLICANT: Garcia-Rodriguez, Consuelo  
 ; TITLE OF INVENTION: Primary Sequence and cDNA of  
     ; POSSANI-POSTAY, Lourival D.  
 ; PATENT NO. 6270785  
 ; TITLE OF INVENTION: Insecticidally Effective Toxins from Scorpions of the  
     ; Genus Centruroides  
 ; NUMBER OF SEQUENCES: 9  
 ; CORRESPONDENCE ADDRESS:  
     ; ADDRESSEE: Finegan, Henderson, Farabow, Garrett &  
     ; STREET: 1300 I Street, N.W.  
     ; CITY: Washington  
     ; STATE: DC  
     ; ZIP: 20005  
 COMPUTER READABLE FORM:  
     ; MEDIUM TYPE: Floppy disk  
     ; COMPUTER: IBM PC compatible  
     ; OPERATING SYSTEM: PC-DOS/MS-DOS  
     ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
     ; APPLICATION NUMBER: US/09/053,021  
     ; FILING DATE:  
     ; CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
     ; APPLICATION NUMBER: US 6/017,007  
     ; FILING DATE: 30-APR-1996  
 PRIOR APPLICATION DATA:  
     ; APPLICATION NUMBER: US 08/848,261  
     ; FILING DATE: 29-APR-1997  
 ATTORNEY/AGENT INFORMATION:  
     ; NAME: Garrett, Arthur S.  
     ; REGISTRATION NUMBER: 20,338  
 REFERENCE/DOCKET NUMBER: 06899.0001-01000  
 TELECOMMUNICATION INFORMATION:  
     ; TELEPHONE: (202)408-4000  
     ; TELEFAX: (202)408-4400  
     ; INFORMATION FOR SEQ ID NO: 9:  
         ; SEQUENCE CHARACTERISTICS:

GENERAL INFORMATION:  
 APPLICANT: Liang, Jihong  
 APPLICANT: Shah, Dilip M.  
 APPLICANT: Wu, Yonnie S.  
 APPLICANT: Rosenberger, Cindy A.  
 TITLE OF INVENTION: Antifungal Polypeptide and Methods for Controlling Plant Pathogenic Fungi.  
 NUMBER OF SEQUENCES: 19  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Charles E. Cohen, Monsanto Company, BB4F  
 STREET: 700 Chesterfield Village Parkway No. 577369th  
 CITY: St. Louis  
 STATE: Missouri  
 COUNTRY: USA  
 ZIP: 63198  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 REFERENCE/DOCKET NUMBER: US-09/103,489  
 FILING DATE: 24-JUN-1998  
 CLASSIFICATION: 800  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Cohen, Charles E.  
 REGISTRATION NUMBER: 34,565  
 REFERENCE/DOCKET NUMBER: 38-21 (10700)A  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (314) 537-6224  
 TELEFAX: (314) 537-6047  
 INFORMATION FOR SEQ ID NO: 15:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 79 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE: Peptide  
 APPLICATION NUMBER: US/08/627,706  
 FILING DATE:  
 CLASSIFICATION: 436  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Cohen, Charles E.  
 REGISTRATION NUMBER: 34,565  
 REFERENCE/DOCKET NUMBER: 38-21(10700)A  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (314) 537-6224  
 TELEFAX: (314) 537-6047  
 INFORMATION FOR SEQ ID NO: 15:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 79 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE: Peptide  
 \*US-08-627-706-15

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 REFERENCE/DOCKET NUMBER: US-09/103,489  
 FILING DATE: 24-JUN-1998  
 CLASSIFICATION: 800  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Cohen, Charles E.  
 REGISTRATION NUMBER: 34,565  
 REFERENCE/DOCKET NUMBER: 38-21 (10700)A  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (314) 537-6224  
 TELEFAX: (314) 537-6047  
 INFORMATION FOR SEQ ID NO: 15:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 79 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-09-103-489-15

Query Match Score 17.7%; Pred. No. 2.7;  
 Best Local Similarity 27.5%; Mismatches 9; Indels 12; Gaps 4;  
 Matches 19; Conservative 9;  
 Job time: 126 sec

QY 4 IAIIFIVLYAFCLLE-DGIVEAGF-----GCPENAGKCHRHKSIIRR-RGGFCRGTR 52  
 DB 8 ISLILFAALVLFAAEAPTMDARLCERPSTWSGVCGNNNACRNQCRNLERAEGHSCNYV 67

QY 53 FRT-TCVCY 60  
 DB 68 FPAHKCICY 76

Search completed: September 17, 2002, 15:44:35

RESULT 15  
 US-09-103-489-15  
 ; Sequence 15, Application US/09/03489  
 ; Patent No. 6215048  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liang, Jihong  
 ; APPLICANT: Shah, Dilip M.  
 ; APPLICANT: Wu, Yonnie S.  
 ; APPLICANT: Rosenberger, Cindy A.  
 ; TITLE OF INVENTION: Antifungal Polypeptide and Methods for Controlling Plant Pathogenic Fungi  
 ; NUMBER OF SEQUENCES: 20  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Charles E. Cohen, Monsanto Company, BB4F  
 ; STREET: 700 Chesterfield Village Parkway No. 6215048th  
 ; CITY: St. Louis  
 ; STATE: Missouri  
 ; COUNTRY: USA  
 ; ZIP: 63198



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: September 17, 2002, 15:42:09 ; Search time 29.91 Seconds  
(without alignments)  
226,530 Million cell updates/sec

Title: US-09-829-481-4

Perfect score: 339  
Sequence: 1 MKSIAIIIFIVLVAFCILEDGG.....IRRGGFCRGTFRTTCYCYR 61

Scoring table: BIOUM62

Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters:

747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_032802:\*

1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT: \*  
 2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT: \*  
 3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT: \*  
 4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT: \*  
 5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT: \*  
 6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT: \*  
 7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT: \*  
 8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT: \*  
 9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT: \*  
 10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT: \*  
 11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT: \*  
 12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT: \*  
 13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT: \*  
 14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT: \*  
 15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT: \*  
 16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT: \*  
 17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT: \*  
 18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT: \*  
 19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT: \*  
 20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT: \*  
 21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT: \*  
 22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	177	52.2	38 19 AAW66437	Cationic Peptide S
2	177	52.2	38 21 AA91736	Cationic Peptide S
3	133.5	39.4	38 15 AAR51160	Antibacterial pept
4	133.5	39.4	38 19 AAW66436	Cationic Peptide I
5	133.5	39.4	38 21 AA91735	Cationic Peptide I
6	71.5	21.1	43 16 AAR82841	Allomycin A antib
7	70	20.6	75 22 AAE10361	Soybean Gly m2 pro
8	70	20.6	94 10 AAP94261	Antibacterial poly
9	69.5	20.5	43 16 AAR82842	Allomycin B antib
10	68.5	20.2	40 14 AAR30562	Sapecin - an antib
11	68.5	20.2	72 21 AAY44509	Pe Defensin prote

Cationic peptide S	40 19 AAW66435	Cationic peptide S
Plant SDF encoded	40 21 AAY91734	Plant SDF encoded
Plant SDF encoded	77 21 AAB24718	Antibacterial poly
Plant SDF encoded	78 21 AAB24718	Plant SDF encoded
Teneacin. Tenebrion	94 10 AWP90001	Teneacin. Tenebrion
Plant SDF encoded	100 21 AAB24717	Plant SDF encoded
Phormia terraenovae	10 18 AFW26000	Plant SDF encoded
Cationic peptide P	18 21 AAB24967	Plant SDF encoded
Yeast pro-AMF C-te	21 AAB24966	Plant SDF encoded
Arabidopsis thalia	21 AAB24965	Raphanus sativus a
Radish antifungal	81.7 21 AAB24965	Radish antifungal
Scorpion sodium ch	40 11 AFR00720	Scorpion sodium ch
Propionibacterium	40 19 AWW66454	Propionibacterium
Radish antifungal	40 21 AY91753	Radish antifungal
Maize basal endosp	45 15 AY91753	Maize basal endosp
Rabbit beta-8 inter	81 20 AY28848	Rabbit beta-8 inter
Human beta-8 inter	80 18 AYW19280	Human beta-8 inter
Arabidopsis thalia	80 18 AYW19617	Arabidopsis thalia
Arabidopsis thalia	85 22 ABB20070	Arabidopsis thalia
Arabidopsis thalia	54 22 APU52922	Arabidopsis thalia
Arabidopsis thalia	80 22 APU61792	Arabidopsis thalia
Arabidopsis thalia	81 20 AY28848	Arabidopsis thalia
Arabidopsis thalia	81 20 AY27683	Arabidopsis thalia
Arabidopsis thalia	76.8 13 AYR27684	Arabidopsis thalia
Human growth hormo	88 13 AYG10630	Human growth hormo
Arabidopsis thalia	88 21 AYG10630	Arabidopsis thalia
Arabidopsis thalia	88 21 AYG45781	Arabidopsis thalia
Arabidopsis thalia	90 21 AYG45780	Arabidopsis thalia
Arabidopsis thalia	106 21 AYG10638	Arabidopsis thalia
Arabidopsis thalia	106 21 AYG45779	Arabidopsis thalia
Human growth hormo	106 21 AYB88036	Human growth hormo
Arabidopsis thalia	106 21 AYB04820	Arabidopsis thalia
Amino acid sequenc	77 21 AYB3873	Amino acid sequenc
Antifungal protein	77 21 AYB68682	Antifungal protein

## ALIGNMENTS

RESULT 1	AAW66437	AAW66437 standard; peptide; 38 AA.
XX	AAW66437	XX
XX	AAW66437;	AC
XX	AAW66437;	XX
DT 12-JAN-1999 (first entry)	XX	XX
cationic peptide scorpion defensin.	DE	DE
Indolicidin analogue; resistance; cationic peptide; antibiotic; bacterial infection; tolerance; antibacterial; microorganism; bacteria; fungus; parasite; virus.	XX	XX
Leiurus quinquestriatus.	OS	OS
PN WO9840401-A2.	PN	PN
XX	XX	XX
PD 17-SEP-1998.	PD	PD
XX	XX	XX
PF 10-MAR-1998;	PF	PF
98WO-QA00190.	98US-0030619.	98US-0030619.
PR 25-FEB-1998;	PR	PR
98US-0030619.	PR 10-MAR-1997;	97US-0040649.
97US-0040649.	PR 20-AUG-1997;	97US-0015314.
97US-0015314.	PR 26-SEP-1997;	97US-0060099.
97US-0060099.	XX	XX
(MICR-) MICROLOGIX BIOTECH INC.	PA	PA
Fraser JR, McNicol PJ, West MHP;	PI	PI
WPI; 1998-520800/44.	DR	DR
New indolicidin peptide analogues - useful for, e.g. enhancing	PT	PT

activity of antibiotic or overcoming tolerance, acquired resistance or inherent resistance of microorganisms

Disclosure: Page 10; 105pp; English.

AAW66393 to AAW66469 represent native cationic peptides from the present invention. The present invention describes compositions and methods for treating infection, especially bacterial infections. The compositions and methods use cationic peptides in combination with an antibiotic agent which are then administered to a patient to enhance activity of the antibiotic agent, to overcome: (a) tolerance; (b) acquired resistance; and (c) inherent resistance. The combinations of antibiotics and cationic peptides can provide synergistic activity against a microorganism that is tolerant, inherently resistant, or acquired resistance to an antibiotic agent. They can be used for killing e.g. bacteria, fungi, parasites and viruses.

	Sequence	38 AA:	Query Match	52.2%	Score 177;	DB 21;	Length 38;
			Best Local Similarity	73.0%	Pred. No. 1.9e-13;		
			Matches	27;	Mismatches	3;	
X			GFGCPFNAGCHRICKSTERRGGCGRGTRPTTCYR	61			
X			1 gfgcp1nqachrhcrsirrrggycagfkqktctcyr	37			
Q							

Query	Match	Score	Length
Best Local Matches	52.2%; 27; Conservative	177; Pred. No. 1	38;
Similarity	73.0%;	DB 19; 9e-13	
		3; Mismatches	Indels
25	GFGCPFNAGKCHRHKCSIRRGGFCRTRFTCVCYR	61	
1	gfgcpnqaghcnrsirrgggccffkqtctcyr	37	

	Query Match	Score 177;	DB 19;	Length 38;		Accession:	NC
	Best Local Similarity 73.0%	Pred. No. 1.9e-13;				XX	18-OCT-1994 (first entry)
Matches	27; Conservative Matches	3; Mismatches	7; Indels	0; Gaps	0;	DT	
						XX	Antibacterial peptide induced in Aeschna cyanea.
						DE	
						XX	Defensin; antibacterial peptide; Odonata; paleopteran insect;
						KW	dragonfly.
						NN	
Y	25	GGCPNAGKCHRCKSIRRGFGCRGTFRTTCVGYR 61					
b	1	ggcpnagkchrckssirrgfgcrgtfrttcvcyr 37					

XX  
 PH Key Location/Qualifiers  
 FT Disulfide-bond 4..26  
 FT Disulfide-bond 11..34  
 FT Disulfide-bond 14..36  
 FT Region 11..21  
 FT /label= "alpha\_helix  
 FT /note= "approximate position"  
 FT Region 24..38  
 FT /label= "beta-pleated\_sheet  
 FT /note= "approximate position"  
 XX  
 PN FF2695392-A.  
 XX  
 PN 11:MAP-1994  
 treatment;  
 carcinoma;  
 bronchus;

PF 04-SEP-1993; 92FR-0010609.  
 XX  
 PR 04-SEP-1992; 92FR-0010609.  
 XX  
 PA (CNRS ) CNRS CENT NAT RECH SCI.  
 XX  
 PI Bulet P, Hettre C, Hoffmann J;  
 XX

Novel pharmaceutical composition containing optionally activated microtumours for treating tumours  
Novel pharmaceutical composition containing optionally activated microtumours for treating tumours  
**(MICR-1) MICROLOGIX BIOTECH INC.**  
Friedland HD, Krieger TJ, Taylor R, Erflie D, Fraser JR, West MHP;  
WPI: 2000-223549/19.

XX New antibacterial peptide(s) from dragonfly - for medical,  
PT veterinary, agricultural and food preservation use  
PT  
XX  
PS Claim 10; Page 17; 25pp; French.  
PS  
XX

Disclosure; Page 11; 94pp; English.

This sequence represents a cationic peptide amino acid sequence, which can be used in the pharmaceutical composition of the invention. The invention relates to a pharmaceutical composition containing at least one activated polyoxalkylene (APO)-modified cationic peptide. The modification of peptides with APO increases their activity against tumour cells, including those with a multi-drug resistant phenotype. The pharmaceutical composition can be used to treat tumours, specifically

CC antiparallel beta-pleated sheet through 2 of the disulphide bonds.  
CC The peptide has antibacterial activity which makes it suitable for  
CC use in the treatment of bacterial infections of eyes and ears and  
CC for oral/dental hygiene and gynaecology.

Query Match      39.4%;      Score 133.5;      DB 15;  
 Best Local Similarity      55.3%;      Pred. No. 1.9e-08;



XX  
PF 28-DEC-1993; 93JP-0350294.  
XX  
PR 28-DEC-1993; 93JP-0350294.  
PA (AMAN ) AMANO PHARM KK.  
XX  
DR WPI; 1995-299564/39.  
N-PSDB; AAT01306.  
XX  
Novel polypeptide derived from *Allomyrina dichotoma* larvae - is physiologically active against bacteria, and may be used in drugs and foodstuffs.  
XX  
PS Claim 1; Page 10; 14pp; Japanese.  
XX  
Two new peptides have been derived from *Allomyrina dichotoma*, allomycin A and allomycin B (AAR81841 and AAR82842). The new peptides have antibacterial activity which makes them useful in the drug and foodstuff industries. The proteins may be prepared using DNA coding for the peptides (AAT01306 and AAT01307) in standard recombinant techniques, this also enables the large scale production of large quantities of the peptides.  
XX  
Sequence 43 AA;  
  
Query Match Score 21.1%; Best Local Similarity 42.9%; Pred. No. 0.28; Length 43; Matches 15; Conservative 2; Mismatches 15; Indels 3; Gaps 1;  
  
Qy 27 GCPFNAGKCHHCKSIRRGGFCRGTRTTCVCYR 61  
Db 12 9faahns1cahclraigrrgcs---rgvcicrr 43  
  
RESULT 7  
AAE10361 ID AAE10361 standard; Protein; 75 AA.  
XX  
AC AAE10361;  
XX  
DT 10-DEC-2001 (first entry)  
XX  
DE Soybean Gly m2 protein from clone s1s1c.pk027.all.  
XX  
KW Soybean; allergen; transgenic plant; P34 protein; Gly m Bd 30K; Gly m 1; soybean vacuolar protein; Gly m IA; Gly m IB; rGly m3; Glycinin G1; alabib; food; infant formula; animal feed; coating; salad oil; syrup; spraying oil; roasting oil; frying oil; cracker; confectionery product; snack food; topping; sauce; batter; breading mixture; baking mix; dough; Gly m2 protein; Clone s1s1c.pk027.all.  
XX  
OS Glycine max.  
XX  
PN WO200168887-A2.  
XX  
PD 20-SEP-2001.  
XX  
PR 15-MAR-2001; 2001WO-US082254.  
XX  
PR 16-MAR-2000; 2000US-0189823.  
XX  
PA (DUPO ) DU PONT DE NEMOURS & CO E I .  
PA (PTON ) PIONEER HI-BRED INT INC.  
XX  
PI Jung R, Kinney AJ;  
XX  
WPI; 2001-582469/65.  
N-PSDB; AAD17530.  
XX  
Recombinant expression construct to lower allergen (e.g. Gly m Bd 30K) content of a soybean, comprises a nucleic acid fragment encoding the  
PT allergen, useful for producing soybean plants which can be used to make soybean products -  
XX  
PS Claim 56; Page 52; 57pp; English.  
XX  
The patent discloses hypoallergenic transgenic soybeans and recombinant expression constructs to lower soybean vacuolar protein, commonly known as P34 (Gly m Bd 30K or Gly m 1) and other allergens such as Gly m IA, Gly m IB, rGly m3 or Glycinin G1 (alabib). The allergen content of the soybean is reduced by sense suppression which is accomplished by using the expression construct that comprises a nucleic acid fragment encoding the allergen. The constructs are useful for producing hypoallergenic transgenic soybean plants which can be used to make hypoallergenic soybean products which can be used in a variety of food (e.g. infant formulas) and animal feed applications. The oil made from seeds of the soybean plants can be used as ingredients, as coatings, as salad oils, as roasting oils, and as frying oils. The foods in which the oil may be used include crackers and snack foods, confectionery products, syrups and toppings, sauces, batter and breading mixtures, baking mixes and doughs. The present sequence is soybean Gly m2 protein from s1s1c.pk027.all clone. This sequence is a minor soybean seed allergen.  
CC  
Sequence 75 AA;  
  
Query Match Score 70; DB 22; Length 75;  
Best Local Similarity 32.4%; Pred. No. 0.73; Gaps 4;  
Matches 23; Conservative 7; Mismatches 23; Indels 18; Gaps 4;  
CC  
Qy 2 KSIAT-LIFIVLVAFCILE-----DGTIVEAGFGCPFMAGKCHRCKS---IRRGG 46  
Db 4 ksiagifcfiflvfvaqevvqteaktcenladtyrgpcfttgsodchcknkehllr--- 60  
CC  
Qy 47 GFCRGTFRTTC 57  
Db 61 grcddfrfcwc 71  
XX  
RESULT 8  
AAP94261 ID AAP94261 standard; protein; 94 AA.  
XX  
AC AAP94261;  
XX  
DT 17-JUN-1990 (first entry)  
XX  
DE Antibacterial polypeptide precursor of *Sarcophaga peregrina*.  
XX  
KW Antibacterial polypeptide precursor; *Sarcophaga peregrina*.  
XX  
OS Sarcophaga peregrina.  
XX  
FH Key Location/Qualifiers  
FT Region 55..94  
FT /note="cloned antibacterial polypeptide"  
FT Region 24..52  
FT /note="previously disclosed sequence"  
FT Region 1..23  
FT /label=signal sequence  
XX  
PN EP303858-A.  
XX  
PD 22-FEB-1989.  
XX  
FT 26-JUL-1988;  
XX  
PR 20-AUG-1987;  
XX  
PA (SANN ) SANWA KAGAKU KENKYUSHO CO. LTD.  
XX  
PI Sawai K, Natori S, Takahashi H, Tanaka K, Mitanai T, Kurono M;  
XX



XX	13-JAN-2000.	PR	26-SEP-1997;	97US-0060099.
XX	02-JUL-1999;	PA	(MICR-) MICROLOGIX BIOTECH INC.	
PF	99WO-CA006008.	XX		
XX		P1	Fraser JR, McNicol PJ, West MHP;	
PR	03-JUL-1998;	XX		
PR	98CA-2242116.	DR	WPI; 1998-520800/44.	
PR	06-JUL-1998;	XX		
XX	98US-0091751.	PT	New indolicidin peptide analogues - useful for, e.g. enhancing	
PA	(YUMA-) UNIV MANITOBA.	PT	activity of antibiotic or overcoming tolerance, acquired resistance	
XX		PT	or inherent resistance of microorganisms	
PI		XX		
PT		PS	Disclosure; Page 10; 105pp; English.	
XX		XX		
WPI:	2000-126938/11.	AAW66393 to AAW66469 represent native cationic peptides from the		
DR	N-PSDB; AAZ29412.	CC	present invention. The present invention describes compositions and	
XX		CC	methods for treating infection, especially bacterial infections. The	
PT	Recombinant expression system for expressing DRR206 or defensin, used	CC	compositions and methods use cationic peptides in combination with an	
PT	to produce pathogen resistant Brassica napus -	CC	antibiotic agent which are then administered to a patient to enhance the	
XX		CC	activity of the antibiotic agent, to overcome: (a) 'tolerance'; (b)	
XX	Claim 11; Fig 9; 39pp; English.	CC	acquired resistance; and (c) inherent resistance. The combinations of	
CC	The present amino acid sequence is the defensin protein, derived from	CC	cationic and cationic peptides can provide synergistic activity	
CC	the clone Drr230 of Pea. This sequence is induced by bacterial and	CC	against a microorganism that is tolerant, inherently resistant, or has	
CC	fungal pathogens like, blackleg fungus. The gene sequence is used in a	CC	acquired resistance to an antibiotic agent. They can be used for killing	
CC	recombinant expression system, capable of transforming plants like,	CC	e.g. bacteria, fungi, parasites and viruses.	
CC	canola (Brassica napus), under the constitutive control of CaMV 35S	XX		
CC	inducible promoter, responsive to pathogen infections. T-DNA sequence is	SQ	Sequence 40 AA;	
CC	also present, for integration of the expression system into the plant			
CC	genome. Transgenic plants expressing DRR206 and defensin protein,			
CC	inhibited fungal growth invitro and are resistant to pathogenic			
CC	infections of Rhizoctonia solani, Leptosphaeria maculans and Sclerotinia			
CC	sclerotiorum.			
XX	Sequence 72 AA;			
Query	Match	20.1%	Score 68;	DB 19;
Best	Local	39.4%	Length 40;	
Matches	13;	Conservative	Pred. No. 0.5;	
Db	Mismatches	2;	Mismatches 16;	Indels 2;
Qy	RESULT	13	Gaps 1;	Gaps 1;
27	AAV91734			
Qy	ID			
2	AAV91734 standard; Peptide; 40 AA.			
4	Db			
ksiacslflvlvfvaedivseantcenlagsykgvcfgcdhert--qegalsgrcr 61	8	gtglnhsacaancllrgnrgyngc-kavcvc 38		
Qy	AC			
51	AAV91734;			
GPFRITC 57	XX			
DE	DT	06-JUN-2000 (first entry)		
62	XX			
ddfrcw	DE	Cationic peptide sapecin amino acid sequence.		
68	XX			
Qy	RESULT	12		
2	AAW66435			
KSTAIIFIVVAFCILEDGVIVAGEGCOPENACK-----CHRACKSIRRG--GFGR 50	XX	Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;		
4	ID			
ksiacslflvlvfvaedivseantcenlagsykgvcfgcdhert--qegalsgrcr 61	XX	KW leukaemia; polyoxalkylene-modified; APO; lymphoma; multiple myeloma;		
Qy	AC			
51	AAV91734;			
GPFRITC 57	XX			
DE	XX			
62	XX			
ddfrcw	DE			
68	XX			
Qy	RESULT	12		
2	AAW66435			
KSTAIIFIVVAFCILEDGVIVAGEGCOPENACK-----CHRACKSIRRG--GFGR 50	XX	Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;		
4	ID			
AAW66435;	XX	KW leukaemia; polyoxalkylene-modified; APO; lymphoma; multiple myeloma;		
XX	AC			
12-JAN-1999	XX			
(first entry)	OS			
XX	DT	Unidentified.		
DE	XX			
Cationic peptide sapecin.	OS			
XX	PN			
Indolicidin analogue; resistance; cationic peptide; antibiotic;	PN			
bacterial infection; tolerance; antibacterial; microorganism;	XX			
XX	PD			
bacteria; fungus; parasite; virus.	XX			
XX	PT			
Sacrophaga peregrina.	XX			
XX	PA			
WO9840401-A2.	XX			
XX	PI			
17-SEP-1998.	XX			
XX	DR			
10-MAR-1998;	XX			
98US-CA001190.	PT			
XX	PR			
25-FEB-1998;	PT			
PR 10-MAR-1997;	PT			
98US-0030619.	PT			
PR 20-AUG-1997;	PT			
97US-0915314.	XX			

PS Disclosure; Page 11; 94pp; English.  
 XX This sequence represents a cationic peptide amino acid sequence, which  
 CC can be used in the pharmaceutical composition of the invention. The  
 CC invention relates to a pharmaceutical composition containing at least one  
 CC activated polyoxalkylene (APO)-modified cationic peptide. The  
 CC modification of Peptides with APO increases their activity against tumour  
 CC pharmaceutical composition can be used to treat tumours, specifically  
 CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,  
 CC cervix, uterus, skin, prostate, liver and colon.  
 XX Sequence 40 AA;

Query Match 20.1%; Score 68; DB 21; Length 77;  
 Best Local Similarity 30.1%; Pred. No. 1 3;  
 Matches 22; Conservative 9; Mismatches 26; Indels 16; Gaps 4;  
 SQ Sequence 77 AA;

Query Match 20.1%; Score 68; DB 21; Length 77;  
 Best Local Similarity 30.1%; Pred. No. 1 3;  
 Matches 22; Conservative 9; Mismatches 26; Indels 16; Gaps 4;  
 SQ Sequence 77 AA;

Query 2 KSIAIIFTIVLVARCILEDGV-EAGFGCPEN-----AGKCHRHKCSIRRRG-G 47  
 Qy | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 3 k1qsvfrwmirtlvgvaneglqkpkqgneilekgsncvaecdsmcvkkrggag 62  
 Qy 48 FCGRTRTTCVY 60  
 Qy | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 63 ycspskk -cycy 73

RESULT 15  
 AAB24718  
 ID AAB24718 standard; Peptide: 78 AA.  
 XX DT 27-NOV-2000 (first entry)  
 AC AAB24718;  
 AC AAB24718;  
 DE Plant SDF encoded polypeptide sequence SEQ List 1 NO:161.  
 XX DE Plant SDF encoded polypeptide sequence SEQ List 1 NO:161.  
 AC AAB24719  
 XX AC AAB24719;  
 XX DT 27-NOV-2000 (first entry)  
 XX DE Plant SDF encoded polypeptide sequence SEQ List 1 NO:162.  
 KW Plant; corn; Arabidopsis thaliana; sequence-determined DNA fragment;  
 KW SDF; genetic mapping; identification; promoter; structural gene; UTR;  
 KW untranslated region; expression control.  
 XX OS  
 XX PN WO200040695-A2.  
 XX PD 13-JUL-2000.  
 XX PF 07-JAN-2000; 2000WO-US00466.  
 XX PR 08-JAN-1999; 99US-0115293.  
 XX PA (CERE-) CERES INC.  
 XX PI Alexandrov N, Brover V, Chen X, Subramanian G, Troukhان МЕ;  
 XX PT New corn plant and Arabidopsis thaliana sequence-determined DNA  
 XX fragments, useful for expressing gene products and for controlling  
 XX expression of a target gene -  
 XX DR 2000-465970/40.  
 XX PS Claim 14: Page 400; 673pp; English.  
 XX The present invention describes polynucleotides, such as complete cDNA  
 CC sequences and/or sequences of genomic DNA encompassing complete genes,  
 CC portions of genes, and/or intergenic regions, collectively referred to  
 CC as sequence-determined DNA fragments (SDFs), from corn plants and  
 CC Arabidopsis thaliana. The SDFs are promoters, structural genes,  
 CC untranslated regions (UTRs), or 3' termination sequences. They can be  
 CC used for expressing a gene product and controlling expression of a  
 CC target gene, either as a promoter, a structural gene, an UTR or as a  
 CC 3' termination sequence. They are also useful as tools for genetic  
 CC mapping, and identification of a particular individual plant or for  
 CC clustering a group of plants with a common trait. AAA78433 to AAA78630  
 CC and AAB24605 to AAB25099 represent the specifically claimed  
 CC polynucleotide sequences and polypeptides encoded by them given in the  
 CC present invention.  
 XX SQ Sequence 78 AA;

Query	Match	20.1%	Score	68;	DB	21;	Length	78;
Best Local Matches	Local Similarity	30.1%	Pred. No.	1.3;				
22;	Conservative Matches	9;	Mismatches	26;	Indels	16;	Gaps	4;
Qy	2	KSIAIIPIVLYAFCTILEDGVY-EAGFCGPFN-					-AGKCHRHCKSIRRRG-	G 47
Ddb	4	klqvstvmliftlivilgvvaneglkpkkqcnleilmqnsncvaecdsmcvkkrgkggq					63	
Qy	48	FCRGTFRTTCVY	60					
Ddb	64	ycpspskk- -cvcy	74					

Search completed: September 17, 2002, 15:44:15  
Job time: 126 sec

Result No.	Score	Query	Match	Length	DB	ID	Description
1	184	P56686 arabidopsis thaliana	DEF4_ANDAU	37	1	DEF4_ANDAU	P56686 arabidopsis thaliana
2	177	P31530 brassica napus	DEF4_LEIOH	38	1	DEF4_LEIOH	P31530 brassica napus
3	133	P30224 centruroides ornatus	DEF1_AEGAE	38	1	DEF1_AEGAE	P30224 centruroides ornatus
4	128	P30223 centruroides ornatus	SCX1_CENNO	38	1	SCX1_CENNO	P30223 centruroides ornatus
5	127	P01486 deinococcus radiodurans	DEF1_MYTGA	38	1	DEF1_MYTGA	P01486 deinococcus radiodurans
6	114	P01486 deinococcus radiodurans	DEF1_MYTID	37	1	DEF1_MYTID	P01486 deinococcus radiodurans
7	82.5	P01486 deinococcus radiodurans	PROTEIN_MYTED	35	1	PROTEIN_MYTED	P01486 deinococcus radiodurans
8	80.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans
9	79.5	P01486 deinococcus radiodurans	DEF1_MYTGA	35	1	DEF1_MYTGA	P01486 deinococcus radiodurans
10	79.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans
11	71.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans
12	71.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans
13	71.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans
14	71	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans
15	69.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans
16	69	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans
17	69	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans
18	69	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans
19	68.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans
20	68.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans
21	68	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans
22	66.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans
23	66.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans
24	66	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans
25	65.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans
26	65.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans
27	64.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans
28	63.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans
29	63.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans
30	63.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans
31	63.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans
32	63.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans
33	63	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans

ALIGNMENTS									
<b>RESULT 1</b>									
DEF4_ANDAU STANDARD; PRT; 37 AA.									
ID DEF4_ANDAU P81618; AC P56686; DT 15-JUL-1999 (Rel. 38, Created); DT 15-JUL-1999 (Rel. 38, Last sequence update); DT 30-MAY-2000 (Rel. 39, Last annotation update); DE 4 kDa defensin.									
OS Androctonus australis hector (Sahara scorpion). OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones; OC Buthoidea; Buthidae; Androctonus. OC NCBI_TaxID=70175; RN [1]									
RP SEQUENCE, AND CHARACTERIZATION. RC TISSUE-Hemolymph; RX MEDLINE=57094646; PubMed=893980; RA Ehret-Sabatier L., Loew D., Goyffon M., Fehlbaum P., Hoffmann J.A., RA van Dorsselaer A., Bulot P.; RT "Characterization of novel cysteine-rich antimicrobial peptides from scorpion blood." RL J. Biol. Chem., 271:29537-29544 (1996). CC -i- FUNCTION: ACTIVE AGAINST GRAM-POSITIVE BACTERIA. CC -i- SUBCELLULAR LOCATION: MW=4206.8; METHOD=Electrospray. CC -i- MASS SPECTROMETRY: MW=4206.8; BELONGS TO THE ARTHROPOD DEFENSIN FAMILY. CC -i- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY. DR InterPro: IPR01542; Arthro_defensin. DR Pfam: PF01097; Arthro_defensin. DR PROSITE: PS00425; ARTHROPOD_DEFENSINS_1. KW Antibiotic. FT DISULFID - 4 25 BY SIMILARITY. FT DISULFID - 11 33 BY SIMILARITY. FT DISULFID - 15 35 BY SIMILARITY. SQ SEQUENCE 37 AA: 4212 MW: AB1263EC3FB84C1 CRC64; Query Match 54.3%; Score 184; DB 1; Length 37; Best Local Similarity 75.7%; Pred. No. 2.2e-14; Matches 28; Conservative 3; Mismatches 6; Indels 0; Gaps 0;									
Database : SwissProt_40_*									
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.									
SUMMARIES									
Result No.	Score	Query	Match	Length	DB	ID	Description		
1	184	P56686 arabidopsis thaliana	DEF4_ANDAU	37	1	DEF4_ANDAU	P56686 arabidopsis thaliana	Q22866 arabidopsis thaliana	O22866 arabidopsis thaliana
2	177	P31530 brassica napus	DEF4_LEIOH	38	1	DEF4_LEIOH	P31530 brassica napus	Q39182 brassica napus	Q39182 brassica napus
3	133	P30224 centruroides ornatus	DEF1_AEGAE	38	1	DEF1_AEGAE	P30224 centruroides ornatus	P15223 centruroides ornatus	P15223 centruroides ornatus
4	128	P30223 centruroides ornatus	SCX1_CENNO	38	1	SCX1_CENNO	P30223 centruroides ornatus	P31530 centruroides ornatus	P31530 centruroides ornatus
5	127	P01486 deinococcus radiodurans	DEF1_MYTGA	38	1	DEF1_MYTGA	P01486 deinococcus radiodurans	P30224 deinococcus radiodurans	P30224 deinococcus radiodurans
6	114	P01486 deinococcus radiodurans	DEF1_MYTID	37	1	DEF1_MYTID	P01486 deinococcus radiodurans	P30224 deinococcus radiodurans	P30224 deinococcus radiodurans
7	82.5	P01486 deinococcus radiodurans	PROTEIN_MYTED	35	1	PROTEIN_MYTED	P01486 deinococcus radiodurans	P30224 deinococcus radiodurans	P30224 deinococcus radiodurans
8	80.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans	P30224 deinococcus radiodurans	P30224 deinococcus radiodurans
9	79.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans	P30224 deinococcus radiodurans	P30224 deinococcus radiodurans
10	79.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans	P30224 deinococcus radiodurans	P30224 deinococcus radiodurans
11	71.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans	P30224 deinococcus radiodurans	P30224 deinococcus radiodurans
12	71.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans	P30224 deinococcus radiodurans	P30224 deinococcus radiodurans
13	71.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans	P30224 deinococcus radiodurans	P30224 deinococcus radiodurans
14	71	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans	P30224 deinococcus radiodurans	P30224 deinococcus radiodurans
15	69.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans	P30224 deinococcus radiodurans	P30224 deinococcus radiodurans
16	69	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans	P30224 deinococcus radiodurans	P30224 deinococcus radiodurans
17	69	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans	P30224 deinococcus radiodurans	P30224 deinococcus radiodurans
18	69	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans	P30224 deinococcus radiodurans	P30224 deinococcus radiodurans
19	68.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans	P30224 deinococcus radiodurans	P30224 deinococcus radiodurans
20	68.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans	P30224 deinococcus radiodurans	P30224 deinococcus radiodurans
21	68	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans	P30224 deinococcus radiodurans	P30224 deinococcus radiodurans
22	66.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans	P30224 deinococcus radiodurans	P30224 deinococcus radiodurans
23	66.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans	P30224 deinococcus radiodurans	P30224 deinococcus radiodurans
24	66	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans	P30224 deinococcus radiodurans	P30224 deinococcus radiodurans
25	65.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans	P30224 deinococcus radiodurans	P30224 deinococcus radiodurans
26	65.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans	P30224 deinococcus radiodurans	P30224 deinococcus radiodurans
27	64.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans	P30224 deinococcus radiodurans	P30224 deinococcus radiodurans
28	63.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans	P30224 deinococcus radiodurans	P30224 deinococcus radiodurans
29	63.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans	P30224 deinococcus radiodurans	P30224 deinococcus radiodurans
30	63	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans	P30224 deinococcus radiodurans	P30224 deinococcus radiodurans
31	63.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans	P30224 deinococcus radiodurans	P30224 deinococcus radiodurans
32	63.5	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans	P30224 deinococcus radiodurans	P30224 deinococcus radiodurans
33	63	P01486 deinococcus radiodurans	DEF1_MYTID	35	1	DEF1_MYTID	P01486 deinococcus radiodurans	P30224 deinococcus radiodurans	P30224 deinococcus radiodurans





01-NOV-1991 (Rel.)	20.	Created
01-NOV-1991 (Rel.)	20.	Last sequence update
01-MAR-1992 (Rel.)	21.	Last annotation update
Zophobas atratus		
DS		
SC		
CC		
DC		
NCBI_TaxID=7074;		
[1]		
SEQUENCE.		
TISSUE_Hemolymph;		
EX MEDLINE=92105112; PubMed=1761552;		
Bulet P., Coccianni S., D'Imareci J.-L., Lambert J., Reichhart J.-M., Hoffmann D., Heitru C., Miranda F., Sampieri C., Smith L.A.; "Insect immunity. Isolation from coleopteran insect of a novel inducible antibacterial peptide and of new members of the insect defensin family." J. Biol. Chem. 266:24520-24525(1991).		
-!- FUNCTION: INVOLVED IN ANTI GRAM-POSITIVE ACTIVITY OF IMMUNE HEMOLYMPH OF Z. ATRATUS.		
-!- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.		
DR PIR; B41711; B41711; C41711.		
DR PIR; C41711; C41711.		
HSSP; P10891; IICAA.		
InterPro; IPR001542; Arthro_defensin.		
InterPro; IPR003614; KnotI.		
Pfam; PF010987; Arthro_defensin_1.		
PRINTS; PRO00271; DEFENSIN.		
SMART; SM00505; KnotI_1.		
PROSITE; PS00425; ARTHROPOD_DEFENSINS_1.		
DR INSECT immunity; Antibiotic.		
DISULFID 3 34 BY SIMILARITY.		
DISULFID 20 40 BY SIMILARITY.		
DISULFID 24 42 BY SIMILARITY.		
VARIANT 30 30 R -> T (IN ISOFORM C).		
SEQUENCE 43 AA: 4453 MW: D63E7B10987FD2F CRC64;		
Query Match 23.7%; Score 80.5%; DB 1; Length 43;		
Best Local Similarity 39.5%; Pred. No. 0.0088;		
Matches 15; Conservative 6; Mismatches 12; Indels 5; Gaps		
25 GF--GCPFNAGKCHRCKSKIRRGGFCRGTFRTTCVC 59		
26     :     :     :     :     :		
7 GFEIAGTKLNAAAGHCLALGRGGYCNS - KSVVCYC 42		
RESULT 9		
SCX2_ANDAU STANDARD PRT; 85 AA.		
AC SCX2_ANDAU		
DD P01484;		
DT 21-JUL-1986 (Rel.) 01. Created		
DT 01-JUL-1993 (Rel.) 26. Last sequence update		
DT 16-OCT-2001 (Rel.) 40. Last annotation update		
DE Neurotoxin II precursor (Aah II).		
DS Androctonus australis hector (Sahara scorpion).		
DC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones; Butohidae; Buthidae; Androctonus.		
NCBI_TaxID=70175;		
[1]		
SEQUENCE FROM N.A.		
RRR MEDLINE=90037062; PubMed=2808423;		
RRR Bougis P.E., Rochat H., Smith L.A.;		
RRR "Precursors of Androctonus australis scorpion neurotoxins. Structure of precursors, processing outcomes, and expression of a functional recombinant toxin II."		
RRR J. Biol. Chem. 264:19259-19265(1989).		
[2]		
SEQUENCE OF 20-82.		
RRR MEDLINE=73025153; PubMed=4342910;		
RRR Rochat H., Rochat C., Sampieri F., Miranda F., Lissitzky S.;"The amino-acid sequence of neurotoxin II of Androctonus australis		



ace 6

CC SUBUNIT: MONOMER (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIALTY: EXPRESSED IN FLOWERS AND IN YOUNG FRUITS.

CC -!- DEVELOPMENTAL STAGE: ACCUMULATE DURING RIPENING.

CC -!- SIMILARITY: BELONGS TO THE GAMMA-PUROTHIONIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: X95730; CAA65046\_1; -.

CC HSSP: P20230; 1GPT; DR InterPro; IPR002118; Gamma\_thionin.

CC DR InterPro; IPR003614; Knot1.

CC DR Pfam; PF00304; Gamma-thionin; 1.

CC DR PRODOM; PD002594; Gamma-thionin; 1.

CC SMART; SMD0505; Knot1; 1.

CC DR PROSITE; PS00940; GAMMA\_THIONIN; 1.

CC DR PROSITE; PS00940; GAMMA\_THIONIN; 1.

KW Defensin; Fungicide; Signal.

CC SIGNAL 1 27 BY SIMILARITY.

CC FT CHAIN 28 74 DEFENSIN J-1.2.

CC FT DISULFID 30 74 BY SIMILARITY.

CC FT DISULFID 41 61 BY SIMILARITY.

CC FT DISULFID 47 68 BY SIMILARITY.

CC FT DISULFID 51 70 BY SIMILARITY.

CC SEQUENCE 74 AA; 8249 MW; D92D8F06F3E1552 CRC34;

Query Match Score 71.5%; DB 1; Length 74;

Best Local Similarity 30.0%; Fred. No. 0.13;

Matches 21; Conservative 8; Mismatches 28; Indels 13; Gaps 3;

CC Qy 2 KSIATIIFTIVLVAFCLIEDGIVEAGFGCP-----FNAGKCHRHKCSIRRGGFCRG 51

CC Db 6 KVIAFFLMAMM-LWFATGMVAEARCTCESOSHFRFKGLCFSKNSNCGVCHTEGFNGGHCRG 63

RESULT 14

CC DEFIC\_AEDAE STANDARD; PRT; 99 AA.

CC DEFIC\_AEDAE STANDARD; PRT; 99 AA.

CC P81609; Q9Y0FL1; Q9Y0R0; MEDLINE=991243699; PubMed=9927179; DT 30-MAY-2000 (Rel. 39, Created)

CC NCBI\_TAXID=7159; RN [1] DE

CC Pterygota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Aedes aegypti; Neoptera; Endopterygota; Diptera; Nematocera; Culicidae; Aedes aegypti.; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicidae; Aedes aegypti.

CC SEQUENCE FROM N.A., DEVELOPMENTAL STAGE, INDUCTION, AND SUBCELLULAR LOCATION.

CC STRAIN=LIVERPOOL; TISSUE=fat body; MEDLINE=951243699; PubMed=9927179; DT 16-OCT-2001 (Rel. 40, Last sequence update)

CC RA Lowenberger C.A.; Smartt C.T.; Bulet P.; Ferding M.T.; Severson D.W.; Hoffmann J.A.; Christensen B.M.; De

CC Teneacin 1 precursor. "Insect immunity: molecular cloning, expression, and characterization of cDNAs and genomic DNA encoding three isoforms of insect defensin in Aedes aegypti." Insect Mol. Biol. 8:107-118 (1999).

CC RN [2] RP SEQUENCE OF 60-99, AND INDUCTION.

CC RC STRAIN=LIVERPOOL; MEDLINE=953360030; PubMed=7633471; DT 01-MAR-2002 (Rel. 41, Last annotation update)

CC RA Lowenberger C., Bulet P., Charlet M., Retru C., Hodgeman B., Christensen B.M., Hoffmann J.A.

CC RT "Insect immunity: isolation of three novel inducible bacterial defensins from the vector mosquito, *Aedes aegypti*." Insect Biochem. Mol. Biol. 25:867-873 (1995).

CC RL J. Biochem. 116:53-58 (1994).

CC RX MEDLINE=95096025; PubMed=7798186;

CC RA Moon H.-J.; Lee S.-Y.; Kurata S.; Natori S.; Lee B.-L.; Tenebrionidae; Tenebrio; Tenebrio molitor.

CC RT "Purification and molecular cloning of cDNA for an inducible anti-bacterial protein from larvae of the coleopteran, *Tenebrio molitor*." J. Biochem. 116:53-58 (1994).

CC RT FUNCTION: BACTERICIDAL PROTEIN PRODUCED IN RESPONSE TO INJURY. IT IS CYTOTOXIC TO GRAM-POSITIVE BACTERIA.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.

[3]

SEQUENCE OF 60-99.  
STRAIN=REFM; PubMed=1568225;  
MEDLINE=9607965; Pubmed=1568225;

RC Chalk R. Albuquerque C.M., Ham P.J., Townson H.;  
RA "Full sequence and characterization of two insect defensins: immune peptides from the mosquito Aedes aegypti.";  
RT Peptides. R. Soc. Lond. B, Biol. Sci. 261:217-221(1995).  
RL Proc. R. Soc. Lond. B, Biol. Sci. 261:217-221(1995).

CC -1- FUNCTION: ANTIBACTERIAL PEPTIDE MOSTLY AGAINST GRAM-POSITIVE BACTERIA.  
CC -1- SUBCELLULAR LOCATION: SECRETED IN THE HEMOLYMPH.  
CC -1- DEVELOPMENTAL STAGE: EXPRESSED 30 MINUTES AFTER INFECTION AND REMAINED PRESENT THROUGH TO 21 DAYS. EXPRESSED IN WHITE OR CALLOWE PUPAE DURING METAMORPHOSIS, BUT NO EXPRESSION WAS SEEN IN LARVAE.  
CC -1- INDUCTION: By bacterial infection.  
CC -1- MORPHISM: THERE ARE TWO DEFENSIN C ISOFORMS, C1 (SHOWN HERE) AND C2.  
CC -1- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR EMBL; AF156092; AAD40116.2;  
DR EMBL; AF156093; AAD40117.2; .  
DR HSSP; P10891; IICCA.  
DR InterPro; IPRO01542; Arthro\_defensin.  
DR InterPro; IPRO03614; Knot1.  
DR PRINTS; PRO01097; Arthro\_defensin; 2.  
DR SMART; SM00271; DEFENSIN.  
. DR PROSITE; PS00425; Arthropod\_DEFENSINS\_1.  
DR Insect\_immmunity; Antibiotoc; Defensin; Signal; Polymorphism.  
FT SIGNAL 1 23 POTENTIAL.  
FT PROTEP 24 59 DEFENSIN\_C.  
FT CHAIN 60 99 DEFENSIN C.  
FT DISULFID 62 89 BY SIMILARITY.  
FT DISULFID 75 95 BY SIMILARITY.  
FT DISULFID 79 97 BY SIMILARITY.  
FT VARIANT 26 26 G -> E (IN ISOFORM C2).  
FT VARIANT 34 34 P -> S (IN ISOFORM C2).  
FT CONFLICT 83 83 R -> G (IN REF. 2 AND 3).  
FT CONFLICT 91 91 A -> S (IN REF. 2 AND 3).  
SQ SEQUENCE 99 AA; 10696 MW; 7C3566A8F4548ED CRC64;

SEQUENCE OF 60-99.  
STRAIN=REFM; PubMed=1568225;  
MEDLINE=9607965; Pubmed=1568225;

RC Urdangarin M.C., de la Canal L.;  
RA "A defensin gene expressed in sunflower inflorescence.";  
RT Plant Physiol. Biochem. 38:253-258(2000).  
RL CC -1- FUNCTION: MAY PLAY A PROTECTIVE ROLE IN FLOWERS BY PROTECTING THE REPRODUCTIVE ORGANS FROM POTENTIAL PATHOGEN ATTACK.  
CC -1- SUBCELLULAR LOCATION: CELL WALL OR VACUOLE (POTENTIAL).  
CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN FLOWERS AND TO A LESSER EXTENT IN LEAVES. LOWER LEVELS IN HYPOCOTYLUS. NO EXPRESSION IN ROOTS AND COOTYLLEONS.  
CC -1- DEVELOPMENTAL STAGE: EXPRESSED PROGRESSIVELY DURING FLOWER DEVELOPMENT REACHING THE HIGHEST LEVEL IN THE MATURE FERTILIZED FLOWER STAGE.  
CC -1- SIMILARITY: BELONGS TO THE GAMMA-PYROTHIONIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR EMBL; AF178634; AAF72042.1; .  
DR EMBL; AF141131; AAF6591.1; .  
DR InterPro; IPRO02118; Gamma-thionin.  
DR InterPro; IPRO03614; Gamma-thionin.  
DR Pfam; PF00304; Gamma-thionin; 1.  
DR PRINTS; PRO00288; PYROTHIONIN.  
DR Frodon; PD002594; Gamma-thionin; 1.  
DR SMART; SM00505; Knot1; 1.  
DR PROSITE; PS00940; GAMMA\_THIONIN; 1.  
KW Thionin; Plant toxin; Signal.  
FT SIGNAL 1 20 POTENTIAL.  
FT CHAIN 21 78 FLOWER SPECIFIC GAMMA-THIONIN.  
FT DISULFID 34 78 BY SIMILARITY.  
FT DISULFID 45 65 BY SIMILARITY.  
FT DISULFID 51 72 BY SIMILARITY.  
FT DISULFID 55 74 BY SIMILARITY.  
SQ SEQUENCE 78 AA; 8627 MW; IAD6A98582149A2D CRC64;

Query Match Best Local Similarity 36.0%; Score 70.5%; DB 1; Length 78;  
Matches 27; Conservative 6; Mismatches 25; Indels 17; Gaps 5;

QY 1 MKSIAIIF--TVLVAFCL--EDG--IVEAAGF-----GCPFNAGKCHRICKSIRR 44  
DB 1 MKSSMMRFAALLVVVCLLANEMGGELVARTCESQSHFKFGTCLSDTNCANVCHSERF 60

Query Match Best Local Similarity 41.7%; Score 71; DB 1; Length 99;  
Matches 15; Conservative 5; Mismatches 12; Indels 4; Gaps 2;

QY 24 AGFGCPFNAGKCHRICKSIRRGGFCRGTCVTCVC 59  
DB 66 SGFGYGDSSA--CAAHCIAFRNRRGYCNA-KKVVCVC 97

RESULT 15  
THGF\_HELAN STANDARD PRT; 78 AA.  
ID THGF\_HELAN  
AC P82659;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Flower-specific gamma-thionin precursor (Defensin SD2).  
GN Helianthus annuus (Common sunflower).  
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; euasterids II; Asterales; Asteroidae;  
OC Heliantheae; Helianthus.  
NCBI\_TaxID=4232;

Search completed: September 17, 2002, 15:45:49  
Job time: 125 sec



Result No.	Score	Query	Match	Length	DB	ID	Description
1	177	52.2	38	2	JN0613		defensin 4 K - scor
2	133.5	39.4	38	2	S27242		defensin - blue da
3	128	37.8	38	2	S74086		defensin - Mediter
4	80.5	23.7	43	2	B41711		defensin B - beetl
5	79.5	23.5	85	1	NTSR2A		neurotoxin II prec
6	74.5	22.0	43	2	C41711		defensin C - beetl
7	72.5	21.4	43	2	JC2554		holotrichin 1 - Hol
8	71.5	21.1	84	2	JX0332		tenecin 1 precursor
9	69.5	20.5	65	1	NTSRB		neurotoxin XI - sc
10	69	20.4	75	2	S11156		psAS10 protein - c
11	68.5	20.2	72	2	T06599		disease resistance
12	68	20.1	94	2	A31792		sapecin precursor
13	67.5	19.9	79	2	T06381		protease inhibitor
14	66.5	19.6	64	1	NTSRB		neurotoxin III - s
15	66.5	19.6	86	2	A10671		Na+-channel blocki
16	66	19.5	94	2	S12558		defensin A precurs
17	65.5	19.3	64	1	NTSR5L		neurotoxin V - Egy
18	65.5	19.3	64	1	NTSR5M		neurotoxin V - sco
19	64.5	19.0	798	2	A40526		integrin beta-7 ch
20	63.5	18.7	77	2	B84433		protease inhibitor
21	63.5	18.7	80	2	T10823		antifungal protein
22	63.5	18.7	768	2	B41029		integrin beta-8 ch
23	63.5	18.7	769	2	A41029		integrin beta-8 ch
24	63	18.6	90	2	B84865		probable trypsin i
25	62	18.3	990	2	151618		nucleolar phosphop
26	61.5	18.1	77	2	S30578		protease inhibitor
27	61.5	18.1	80	2	T10176		antifungal protein
28	61.5	18.1	154	2	T17816		hypothetical prote
29	61	18.0	40	2	JU0222		sapecin C - flesh

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2002, 15:42:49 ; Search time 18.35 Seconds  
(without alignments)  
319.425 Million cell updates/sec

Title: US-09-829-481-4  
Perfect score: 339  
Sequence: 1 MKSIAIIIVLVAFCLEDG.....IRRRGGFCRGTFRTTCVYR 61

Scoring table: BL0SUMG2  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : PIR\_71:  
1: Pirl:\*2: Pir2:\*3: Pir3:\*4: Pirl:\*A: Reference number: JN0613 ; MUID: 93326112  
A: Molecule type: protein  
A: Residues: 1-38 <COC>  
A: Note: this protein is similar to scorpion toxins and insect defensins

#### ALIGNMENTS

RESULT	1	2	3	Mismatches	Indels	0;	Gaps	0;
JN0613	defensin 4K - scorpion (Leiurus quinquestriatus)							
N: Alternative names: antibacterial 4K peptide								
C: Species: Leiurus quinquestriatus hebraeus								
C: Accession: JN0613								
R: Cocianich, S.; Goffon, M.; Bontems, F.; Bouet, P.; Menez, A.; Hoffmann								
Biochem. Biophys. Res. Commun. 194, 17-22, 1993								
A: Title: Purification and characterization of a scorpion defensin, a 4kDa antibacteri								
A: Reference number: JN0613 ; MUID: 93326112								
A: Molecule type: protein								
A: Residues: 1-38 <COC>								
Q: Note: this protein is similar to scorpion toxins and insect defensins								
Query Match 52.2%; Score 177; DB 2; Length 38;								
Best Local Similarity 73.0%; Pred. No. 5.4e-13;								
Matches 27; Conservative 3; Mismatches 7;								
RESULT	25	GFGCPFNAGKCHRCKSIRRGGFCRGTFRTTCVYR 61						
Db	1	GFGCPLNQAGCHRCKSIRRGGYCAGFFKQTCR 37						
Q: Note: this protein is similar to scorpion toxins and insect defensins								
Query Match 52.2%; Score 177; DB 2; Length 38;								
Best Local Similarity 73.0%; Pred. No. 5.4e-13;								
Matches 27; Conservative 3; Mismatches 7;								
RESULT	25	GFGCPFNAGKCHRCKSIRRGGFCRGTFRTTCVYR 61						
Db	1	GFGCPLNQAGCHRCKSIRRGGYCAGFFKQTCR 37						
Q: Note: this protein is similar to scorpion toxins and insect defensins								
Query Match 52.2%; Score 177; DB 2; Length 38;								
Best Local Similarity 73.0%; Pred. No. 5.4e-13;								
Matches 27; Conservative 3; Mismatches 7;								
RESULT	25	GFGCPFNAGKCHRCKSIRRGGFCRGTFRTTCVYR 61						
Db	1	GFGCPLNQAGCHRCKSIRRGGYCAGFFKQTCR 38						

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

8

Score

Match

Length

DB

ID

Description

S74088 - Mediterranean mussel provincialis (Mediterranean mussel)  
 C;Species: *Mitilus galloprovincialis* (Mediterranean mussel)  
 C;Date: 11-Mar-1998 #sequence\_revision 17-Apr-1998 #text\_change 04-Feb-2000  
 C;Accession: S74088  
 R;Hubert, F.; Noel, T.; Roch, P.  
 Bur. J. Biochem. 240, 302-306, 1996  
 A;Title: A member of the arthropod defensin family from edible Mediterranean mussels (Mytilus galloprovincialis)  
 A;Reference number: S74088; MUID:97025319  
 A;Molecule type: protein  
 A;Residues: 1-38 <BUP>  
 A;Experimental source: hemolymph  
 C;Keywords: antibiotic  
 F;4-25,10-33,14-35,21-38/Disulfide bonds: #status predicted

Query Match Best Local Similarity 37.8%; Score 128; DB 2; Length 38;  
 Matches 24; Conservative 2; Mismatches 10; Indels 2; Gaps 2;  
 Oy 25 GFGCPNAGKGRHRCKSTRN-GGFRTGTRTCYCYR 61  
 Db 1 GFGCP-NNYQCHRHKCKSTPGRCGGYCGGXHRLRCTCYR 37

RESULT 4  
 B41711  
 defensin B - beetle (Zophobas atratus)  
 C;Species: Zophobas atratus  
 C;Date: 24-Jul-1992 #sequence\_revision 24-Jul-1992 #text\_change 03-May-1996  
 C;Accession: B41711  
 R;Bulet, P.; Cocianich, S.; Dimarco, J.L.; Lambert, J.; Reichhart, J.M.; Hoffmann, D.; J. Biol. Chem. 266, 24520-24525, 1991  
 A;Title: Insect immunity. Isolation from a coleopteran insect of a novel inducible antibacterial peptide.  
 A;Reference number: A41711; MUID:92105112  
 A;Accession: B41711  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 1-43 <BUU>  
 C;Superfamily: sapecin precursor

Query Match Best Local Similarity 39.3%; Score 80.5%; DB 2; Length 43;  
 Matches 15; Conservative 6; Mismatches 12; Indels 5; Gaps 2;  
 Oy 25 GF---GCPFNAGKGRHRCKSTRRGFCRGTFRTCV 59  
 Db 7 GFEIAGTKLNAAACGAHCLALGRGGYCNS-KSVCVC 42

RESULT 5  
 NTSR2A  
 Neurotoxin II precursor [validated] - Sahara scorpion  
 N;Alternate names: mammalian neurotoxin (Clone pCD402)  
 C;Species: Androctonus australis (Sahara scorpion)  
 C;Date: 24-Apr-1984 #sequence\_revision 12-Apr-1996 #text\_change 15-Sep-2000  
 C;Accession: D34444; A01744  
 R;Boujat, P.E.; Rochat, H.; Smith, L.A.  
 J. Biol. Chem. 264, 19259-19265, 1989  
 A;Title: Precursors of Androctonus australis scorpion neurotoxins. Structures of precursors  
 A;Reference number: A34444; MUID:90037062  
 B;Accession: D34444  
 A;Molecule type: mRNA  
 A;Residues: 1-85 <BOU>  
 B;Cross-references: GB:MB2704; PID:9161144; PID:AAA29949.1; PID:g161145; GB:J05102  
 A;Experimental source: clone pCD402  
 R;Rochat, H.; Rochat, C.; Samper, F.; Miranda, F.; Lissitzky, S.  
 Bur. J. Biochem. 28, 381-388, 1972  
 A;Title: The amino-acid sequence of neurotoxin II of Androctonus australis Hector.  
 A;Accession: A01744

A;Molecule type: protein  
 A;Residues: 20-83 <ROC>  
 R;Kopeyan, C.; Martinez, G.; Lissitzky, S.; Miranda, F.; Rochat, H.  
 Eur. J. Biochem. 47, 483-489, 1974  
 A;Title: Disulfide bonds of toxin II of the scorpion Androctonus australis Hector.  
 A;Reference number: A91228; MUID:75057756  
 A;Contents: annotation; disulfide bonds  
 R;Houset, D.; Habersetzer-Rochat, C.; Astier, J.P.; Fontecilla-Camps, J.C.  
 J. Mol. Biol. 238, 88-103, 1994  
 A;Title: Crystal structure of toxin II from the scorpion Androctonus australis hector  
 A;Reference number: A58108; MUID:94194522  
 A;Contents: annotation; X-ray crystallography, 1.3 angstroms, residues 20-83  
 R;Fontecilla-Camps, J.C.; Houset, D.  
 Submitted to the Brookhaven Protein Data Bank, September 1994  
 A;Reference number: A67142; pDB:1LPTX  
 A;Contents: annotation; X-ray crystallography, 1.3 angstroms, residues 20-83  
 C;Superfamily: scorpion neurotoxin  
 C;Keywords: amidated carboxyl end; neurotoxin; venom  
 F;1-19/Domain: signal sequence #status Predicted <SIG>  
 F;20-83/Product: neurotoxin II #status experimental <MAT>  
 F;31-82,35-55,41-65,45-67/Disulfide bonds: #status experimental  
 F;83/Modified site: amidated carboxyl end (His) (amide in mature form from following  
 Query Match Best Local Similarity 23.5%; Score 79.5; DB 1; Length 85;  
 Matches 16; Conservative 16; Mismatches 27; Indels 7; Gaps 3;  
 Qy 1 MKSIAITIFIVLVAFCILEDGIVEAGFGCPFNAGK --CHRHCKSIRRGGFCR - GTFR 55  
 Db 6 MISLAFLFVGVE--SVKDGIVDDVNCTYFCGRNAYCNEECKLKGESGYCQWASPYGN 63  
 Qy 56 TCVCYR 61  
 Db 64 ACYCYK 69

RESULT 6  
 C41711  
 defensin C - beetle (Zophobas atratus)  
 C;Species: Zophobas atratus  
 C;Date: 24-Jul-1992 #sequence\_revision 24-Jul-1992 #text\_change 03-May-1996  
 C;Accession: C41711  
 R;Bulet, P.; Cocianich, S.; Dimarco, J.L.; Lambert, J.; Reichhart, J.M.; Hoffmann, D.; J. Biol. Chem. 266, 24520-24525, 1991  
 A;Matches: 13; Conserv 14; Mismatches 6; Indels 13; Gaps 5;  
 A;Title: Insect immunity. Isolation from a coleopteran insect of a novel inducible an  
 A;Reference number: A41711; MUID:92105112  
 A;Accession: C41711  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 1-43 <BUU>  
 C;Superfamily: sapecin precursor

Query Match Best Local Similarity 36.8%; Score 74.5%; DB 2; Length 43;  
 Matches 14; Conservative 14; Mismatches 6; Indels 13; Gaps 5;

Qy 25 GF---GCPFNAGKGRHRCKSTRRGFCRGTFRTCV 59  
 Db 7 GFEIAGTKLNAAACGAHCLALGRGGYCNS-KSVCVC 42

RESULT 7  
 JC2554  
 holotrichin 1 - Holotrichia diomphalia  
 C;Species: Holotrichia diomphalia  
 C;Date: 04-Sep-1998 #sequence\_revision 04-Sep-1998 #text\_change 17-Mar-1999  
 C;Accession: JC2554  
 R;Lee, S.Y.; Moon, H.-J.; Kawabata, S.; Kurata, S.; Natori, S.; Lee, B.L.  
 Biol. Pharm. Bull. 18, 457-459, 1995  
 A;Title: A sapecin homologue of Holotrichia diomphalia: Purification, sequencing and  
 A;Accession number: JC2554; MUID:96031330

A;Accession: JC2554  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-43 <LEE>  
C;Superfamily: sapecin precursor

Query Match 21.4%; Score 72.5; DB 2; Length 43;  
Best Local Similarity 39.4%; Pred. No. 0.3; Indels 3; Gaps 1;  
Matches 13; Conservative 5; Mismatches 12; Indels 3; Gaps 1;

Qy 27 GCPFNAGKCHRICKSIRRRGGFCRGTFRTTCVC 59  
Db 12 GIAINDSACAAHCLAMRKGGSKK--QGVCV 41

RESULT 8

JX0332 teneclin 1 precursor - yellow mealworm  
C;Species: *Tenebrio molitor* (yellow mealworm)  
C;Accession: JX0332 #sequence\_revision 28-Oct-1994 #text\_change 20-Jun-2000  
C;Date: 28-Oct-1994  
R;Moon, H.J.; Lee, S.Y.; Kurata, S.; Natori, S.; Lee, B.L.  
J;Biochem, 116, 53-58, 1994  
A;Title: Purification and molecular cloning of cDNA for an inducible antibacterial protein  
A;Reference number: JX0332; MUID:95096025  
A;Accession: JX0332  
A;Molecule type: mRNA  
A;Residues: 1-84 <HOO>  
A;Cross-references: GB:D17670; NID:q1235940; PIDN:BA0A4552.1; PID:q1235941  
A;Experimental source: larva  
C;Comment: Teneclin 1 is an antibacterial protein induced in the hemolymph of larvae of *Tenebrio molitor*  
C;Superfamily: sapecin precursor  
F;1-11/Domain: signal sequence #status predicted <SIG>  
F;42-84/Product: teneclin 1 #status predicted <MAT>  
F;44-75,61-81,65-83/Disulfide bonds: #status predicted

Query Match 21.1%; Score 71.5; DB 2; Length 84;  
Best Local Similarity 27.3%; Pred. No. 0.48; Indels 26; Gaps 3;  
Matches 21; Conservative 7; Mismatches 26; Indels 23; Gaps 3;

Qy 4 IAIIFIVLVAFCLIEDGII----VEAG-----FGCPFNAGKCHRICKS 42  
Db 9 VACFTFLQAAFPFLPEAAETIPOGEHIRVKRVTCDIISVEAKVQKINDAACAAHCLFR 68

Qy 43 RRRGGFCRGTFRTTCVC 59  
Db 69 GRSGGYCNG -KRVCVC 83

RESULT 9

NTSREB neurotoxin XI - scorpion (Buthus occitanus)  
C;Species: *Buthus occitanus* tunetanus  
C;Accession: A01746  
C;Accession: A01746  
R;Martin, M.F.; Rochat, H.  
Toxicon 22, 279-291, 1984  
A;Title: Purification of thirteen toxins active on mice from the venom of the North African scorpion *Buthus occitanus* tunetanus  
A;Reference number: A94316; MUID:84224414  
A;Accession: A01746  
A;Molecule type: protein  
A;Residues: 1-65 <MAR>  
C;Superfamily: scorpion neurotoxin  
C;Keywords: neurotoxin; venom  
F;12-63,16-36,22-46,26-48/Disulfide bonds: #status predicted

Query Match 20.5%; Score 69.5; DB 1; Length 65;  
Best Local Similarity 26.0%; Pred. No. 0.66; Mismatches 10; Indels 5; Gaps 2;

Qy 17 LEDGIVEAGFGCPFNAGK --CHRICKSIRRRGGFCR --GTFRTTCVCYR 61  
Db 1 LKDGTYIVDDRNCTYFCGTNAYCNFECVKLKGESYCYQWGRYGNACWCYK 50

RESULT 10

S11156 protein - cowpea  
PSAS10 protein - cowpea  
C;Species: *Vigna unguiculata* (cowpea)  
C;Date: 18-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 09-Jan-2000  
B;Ishibashi, N.; Yamada, D.; Minamikawa, T.  
Plant Mol. Biol. 15, 59-64, 1990  
A;Title: Stored mRNA in cotyledons of *Vigna unguiculata* seeds: nucleotide sequence of  
A;Reference number: S11156; MUID:91355865  
A;Accession: S11156  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-75 <ISH>  
A;Cross-references: EMBL:X16877; NID:q22075; PID:q22076  
C;Superfamily: gamma-thionin

Query Match 20.4%; Score 69; DB 2; Length 75;  
Best Local Similarity 33.3%; Pred. No. 0.84; Mismatches 6; Indels 14; Gaps 4;

Qy 2 KSTIA-TIIFIVLVAFCLIE-----DGIVEAAGFCGPNPAGKCHRICKSIRR-RGGFC 49  
Db 4 KSTAGICFLFLVFLVFAQEVVVQSEARTCENLVDTYRGPCTFTGSCUDDHCKNKEHLLSGRC 63

Qy 50 RGTFRITCVCYR 61  
Db 64 RDDVR - CWCTR 73

RESULT 11

T06599 disease resistance response protein 230 - garden pea  
C;Species: *Pisum sativum* (garden pea)  
C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jun-2000  
C;Accession: T06599  
R;Chiari, C.; Hadwiger, L.A.  
Mol. Plant Microbe Interact. 4, 324-331, 1991  
A;Title: The Fusarium solani-induced expression of a pea gene family encoding high cy  
A;Reference number: 215787; MUID:92190628  
A;Accession: T06599  
A;Status: preliminary; translated from GB/EMBL/DDJB  
A;Molecule type: mRNA  
A;Residues: 1-72 <CHT>  
A;Cross-references: EMBL:L01578; PIDN:AAA79117.1; PID:q169074  
A;Experimental source: cv. Alaska  
C;Genetics:  
A;Gene: DRR230-a  
C;Superfamily: gamma-thionin

Query Match 20.2%; Score 68.5; DB 2; Length 72;  
Best Local Similarity 28.4%; Pred. No. 0.92; Mismatches 11; Indels 13; Gaps 3;

Qy 2 KSTAIIFIVLVAFCLIEDGIVEAAGFCGPFNAGK-----CHRICKSIRRRG---GFCCR 50  
Db 4 KSLACLSFLLVFLVFAOEIVVSANTENLAGSYKGVFGGCDRHCRT - QEGANISGRCR 61

Qy 51 GTFRTTC 57  
Db 62 DDRCWC 68

RESULT 12

A31192 sapecin precursor - flesh fly (*Sarcophaga peregrina*)

N:Alternate names: antibacterial protein  
 C:Species: Sarcophaga peregrina  
 C:Accession: A31792; A31791  
 C:Cross-references: GB-J04053; NID:9161266; PIDN:AAA29984.1; PID:g161267  
 R:Matsuura, K.; Natori, S.  
 J. Biol. Chem., 263, 17112-17116, 1988  
 A;Title: Molecular cloning of cDNA for sapecin and unique expression of the sapecin gene  
 A;Reference number: A31792; MUID:89034216  
 A;Accession: A31792  
 A;Molecule type: mRNA  
 A;Residues: 1-94 <MAT>  
 A;Cross-references: GB-J04053; NID:9161266; PIDN:AAA29984.1; PID:g161267  
 R:Matsuura, K.; Natori, S.  
 J. Biol. Chem., 263, 17112-17116, 1988  
 A;Title: Purification of three antibacterial proteins from the culture medium of NIH-Sar  
 A;Accession: A31791  
 A;Molecule type: protein  
 A;Residues: 55-94 <M92>  
 C:Superfamily: sapecin precursor

Query Match 20.1% Score 68; DB 2; Length 94;  
 Best Local Similarity 39.4%; Pred. No. 1..3; Matches 13; Conservative 2; Mismatches 16; Indels 2; Gaps 1;  
 QY 27 GCPENAGKCHRHCKSIRRGFCRGTFRITCVC 59  
 Db 62 GTINGHSACAAHCLLRGNRGYCNG-KAVCVC 92

RESULT 13 Submitted to the EMBL Data Library, July 1994

A;Description: Isolation and expression of a drought-induced protease inhibitor from soy

A;Reference number: 215640

A;Accession: T06381

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: mRNA

A;Residues: 1-99 <CUS>

A;Cross-references: EMBL:U12150; NID:9533691; PIDN:AAC97524.1; PID:9533692

A;Experimental source: strain Essex; leaf

C:Superfamily: gamma-thionin

Query Match 19.9% Score 67.5; DB 2; Length 79;  
 Best Local Similarity 26.1%; Pred. No. 1..3; Matches 13; Mismatches 27; Indels 11; Gaps 2;  
 QY 1 MKSTAIIFTYLVAFILEDGIVAEAGFGCPFNAGK-----CHRICKSIRRRGGFCR 50  
 Db 8 VSTICVLLLVAVETRMGPTMVAARTCESQSHREKGPCQLSDTNCGSVCRTERFTGSHCR 67

QY 51 GTFRTTCVC 59

Db 68 G-FRRRCFC 75

RESULT 14

NTSR3B Neurotoxin III - scorpion (Butthus occitanus)

C:Species: Butthus occitanus tunetanus

C:Accession: A01745

R;Vargas, O.; Gregoire, J.; Martin, M.F.; Bechis, G.; Rochat, H.

Toxicon 20, 79, 1982

A;Title: Neurotoxins from the venoms of two scorpions: Butthus occitanus tunetanus and Bu

N:Alternate names: antibacterial protein  
 C:Species: Sarcophaga peregrina  
 C:Accession: A01745  
 A;Molecule type: protein  
 A;Residues: 1-64 <VAR>  
 C:Superfamily: scorpion neurotoxin  
 C:Keywords: blocked carboxyl end; neurotoxin; venom  
 F:12-63 16-36 22-46 26-48/Dsulfide bonds; #stats predicted  
 F:64/Modified site: blocked carboxyl end (Asn) (probably amidated) #status experiment

Query Match 19.6% Score 66.5; DB 1; Length 64;  
 Best Local Similarity 22.0%; Pred. No. 1..4;  
 Matches 11; Conservative 12; Mismatches 22; Indels 5; Gaps 2;  
 QY 17 LEDGIVEAGFGCPFNAGK---CHRICKSIRRRGGFCR-GTFRTTCVCYR 61  
 Db 1 VKDGYIVDDRNCTYFCGFRAYCNEECTKLKGESGYCOWASPGNACYC 50

RESULT 15  
 JN0671 Na+-channel-blocking toxin (clone cngtIV) precursor - scorpion (Centruroides noxioides)  
 C:Species: Centruroides noxioides  
 C:Accession: JN0671  
 C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 16-Jul-1999  
 R:Becerril, B.; Vazquez, A.; Garcia, C.; Corona, M.; Bolivar, F.; Possani, L.D.  
 Gene 128, 165-171, 1993  
 A;Title: Cloning and characterization of cDNAs that code for Na+-channel-blocking toxin  
 A;Reference number: JN0669; MUID: 93232983  
 A;Accession: JN0671  
 A;Molecule type: mRNA  
 A;Cross-references: GB-L05062; NID:9304570; PIDN:AAA28287.1; PID:9304571  
 A;Experimental source: venom gland  
 C:Superfamily: scorpion neurotoxin  
 C:Keywords: toxin  
 F:1-197/domain: signal sequence #status predicted <SIG>  
 F:20-86/Product: Na+-channel-blocking toxin (clone cngtIV) #status predicted <NaT>

Query Match 19.6% Score 66.5; DB 2; Length 86;  
 Best Local Similarity 33.3%; Pred. No. 1..7;  
 Matches 23; Conservative 7; Mismatches 26; Indels 13; Gaps 4;  
 QY 1 MKSTAIIT-FTVLVAFILEDGVIEAGFGCPFNAGK-----CHRICKSIRRRGGFCR 50  
 Db 1 MSLLITTAFLCLVLTIGTVWAKDSYLVDVKGCKRKNCYKLGENDYCNRECKMKHRGGSYGYC 60

QY 51 GTFRTTCVC 59  
 Db 61 G---FGCYC 66

Search completed: September 17, 2002, 15:45:00

Job time: 131 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: September 17, 2002, 15:43:24 : Search time 25.63 Seconds  
Perfect score: 339  
Sequence: 1 MKSIAIIIFIVLVAFCTILEDG . . . . . IRRGGFCRGHFRTCVCYR 61  
(Without alignments)  
(411.732 Million cell updates/sec)

Title: US-09-829-481-4  
Scoring table: BL0SUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues  
Total number of hits satisfying chosen parameters: 562222  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_19\_\*  
1: sp\_archaea:  
2: sp\_bacteria:  
3: sp\_fungi:  
4: sp\_invertebrate:  
5: sp\_human:  
6: sp\_invertebrate:  
7: sp\_mammal:  
8: sp\_mhc:  
9: sp\_organelle:  
10: sp\_phage:  
11: sp\_plant:  
12: sp\_rhodent:  
13: sp\_virus:  
14: sp\_vertebrate:  
15: sp\_unclassified:  
16: sp\_virus:  
17: sp\_archeap:  
18: sp\_bacteriop:  
19: sp\_arcpap:  
20: sp\_bacteriop:  
21: sp\_bacteriop:  
22: sp\_bacteriop:  
23: sp\_bacteriop:  
24: sp\_bacteriop:  
25: sp\_bacteriop:  
26: sp\_bacteriop:  
27: sp\_bacteriop:  
28: sp\_bacteriop:  
29: sp\_bacteriop:  
30: sp\_bacteriop:  
31: sp\_bacteriop:  
32: sp\_bacteriop:  
33: sp\_bacteriop:  
34: sp\_bacteriop:  
35: sp\_bacteriop:  
36: sp\_bacteriop:  
37: sp\_bacteriop:  
38: sp\_bacteriop:  
39: sp\_bacteriop:  
40: sp\_bacteriop:  
41: sp\_bacteriop:  
42: sp\_bacteriop:  
43: sp\_bacteriop:  
44: sp\_bacteriop:  
45: sp\_bacteriop:

## ALIGNMENTS

Result No.	Score	Query Match	Length	DB ID	Description
1	154.5	45.6	81	Q9YOA9	Q9yoa9 mytilus gal
2	152	44.8	82	Q9YOA9	Q9yoa9 mytilus gal
3	146.5	43.2	73	Q9BLU3	Q9blj3 ornithodoros
4	140	41.3	73	Q9BLJ4	Q9bj4 ornithodoros
5	108	31.9	57	Q9YOB0	Q9yob0 mytilus gal
6	90	26.5	91	Q9N661	Q9n661 mesobothrus
7	82.5	24.3	85	Q9NCT7	Q9nct7 mesobothrus
8	79.5	23.5	79	Q96049	Q96049 oryctes rhi
9	77	22.7	57	Q9GY06	Q9gyu6 aedes albop
10	75.5	22.3	85	Q9NTB8	Q9nlp8 mesobothrus
11	75.5	22.3	85	Q9GYX2	Q9gyx2 mesobothrus
12	71.5	21.1	85	Q9BLM4	Q9blm4 androctonus
13	70	20.6	102	Q91721	Q91721 anophelis q
14	70	20.6	103	Q9YOB1	Q9yob1 mytilus gal
15	69.5	20.5	72	Q9FPR8	Q9fr81 pisum sativ
16	68.5	20.2	85	Q9GNGB	Q9gng8 mesobothrus

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8

Result No.	Score	Query Match	Length	DB ID	Description
1	154.5	45.6	81	Q9YOA9	Q9yoa9 mytilus gal
2	152	44.8	82	Q9YOA9	Q9yoa9 mytilus gal
3	146.5	43.2	73	Q9BLU3	Q9blj3 ornithodoros
4	140	41.3	73	Q9BLJ4	Q9bj4 ornithodoros
5	108	31.9	57	Q9YOB0	Q9yob0 mytilus gal
6	90	26.5	91	Q9N661	Q9n661 mesobothrus
7	82.5	24.3	85	Q9NCT7	Q9nct7 mesobothrus
8	79.5	23.5	79	Q96049	Q96049 oryctes rhi
9	77	22.7	57	Q9GY06	Q9gyu6 aedes albop
10	75.5	22.3	85	Q9NTB8	Q9nlp8 mesobothrus
11	75.5	22.3	85	Q9GYX2	Q9gyx2 mesobothrus
12	71.5	21.1	85	Q9BLM4	Q9blm4 androctonus
13	70	20.6	102	Q91721	Q91721 anophelis q
14	70	20.6	103	Q9YOB1	Q9yob1 mytilus gal
15	69.5	20.5	72	Q9FPR8	Q9fr81 pisum sativ
16	68.5	20.2	85	Q9GNGB	Q9gng8 mesobothrus

Q39807 glycine max  
Q9qua7 mesobothrus  
Q9p69 mesobothrus  
Q9lvz7 mus musculus  
07717 aedes albop  
P91193 aedes aegyp  
Q9639 aedes aegyp  
Q9xgd9 zea mays (m  
Q9bf52 acalolepta  
Q9f914 citrus para  
Q94874 pyrif  
Q9pk93 chlamydia m  
Q91803 xenopus lae  
Q9j18 mus musculus  
Q9yb8 mesobothrus  
Q84631 paramecium  
P8280 Stomoxys ca  
Q9et61 ratus norv  
Q9iz6 ratus norv  
Q99103 mus musculus  
Q94435 centroioide  
Q9xv87 mesobothrus  
Q95xx6 mesobothrus  
Q45448 caenorhabdi  
Q9py3 homo sapien  
Q0274 homo sapien  
Q9ylu3 mesobothrus  
Q9uac8 mesobothrus  
Q9nzk7 homo sapien

OX NCBI\_TAXID=29158;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=HEMOCYTE;  
RX MEDLINE=20033584; PubMed=10564642;  
RA Mitca G., Vandebulcke F., Hubert F., Roch P.;  
RT Mussel defensins are synthesized and processed in granulocytes then released into the plasma after bacterial challenge. ;  
RL J. Cell Sci. 112:423-424(1999).  
CC -|- MISCELLANEOUS: THE PI OF THIS PROTEIN IS 8.92.  
CC -|- SUBCELLULAR LOCATION: SECRETED.  
CC -|- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.  
CC DR SIGNAL 1 AAD5118.1;  
DR DR InterPro: AF162338; AAD5118.1;  
DR DR InterPro: IPR01542; Arthro\_defensin.  
DR DR InterPro: IPR03614; Knot1.  
DR SMART: SM00505; Knot1. 1.  
DR PROSITE: PS00425; ARTHROPOD\_DEFENSINS; UNKNOWN\_1.  
DR KW Defensin; Antibiotic; Fungicide; Annotation; Signal.  
FT SIGNAL 1 21 POTENTIAL.  
FT PROPEP 22 60 POTENTIAL.  
FT CHAIN 61 81 AMIDATION (POTENTIAL).  
FT MOD\_RES 60 60 AMIDATION (POTENTIAL).  
FT DISULFID 25 46 BY SIMILARITY.  
FT DISULFID 31 54 BY SIMILARITY.  
FT DISULFID 35 56 BY SIMILARITY.  
FT DISULFID 42 59 BY SIMILARITY.  
SQ SEQUENCE 81 AA; 9087 MW;

Query Match	45.68;	Score 154.5;	DB 5;	Length 81;		Best Local Similarity 51.7%;	Pred. No. 1.3e-11;	Mismatches 6;	Indels 10;	Gaps 1;
Best Local Similarity 30%;	Conservative 7;	Mismatches 18;	Indels 3;	Gaps 3;		Matches 30;	Conservative 7;	Mismatches 18;	Indels 3;	Gaps 3;
Qy	5 ATIFIVLVARCILEDGIVEAGFCGPNAKGKCHRCKSIRR-GGGFCRGTRTTCVCY 61					QY	22 VEAGFGCPNAKGKCHRCKSIRR RGGFCRGTRTTCVCY 60			
Db	3 AAFVLLVVGCLIMTD-VATAGGCPNNYA-CHQHCKSTRYGGCAGWFLRCYR 58					Db	34 VRRGYGCPENQYQCHSHSGIRGYKGCKGTFKQTCKY 73			
RESULT	2					RESULT	4			
ID Q9U6U0	PRELIMINARY;	PRT;	82 AA.			ID Q9BLJ4	PRELIMINARY;	PRT;	73 AA.	
AC Q9U6U0;						AC Q9BLJ4;				
DT 01-MAY-2000 (TREMBLrel. 13, Created)						DT 01-JUN-2001 (TREMBLrel. 17, Created)				
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)						DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)				
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)						DE DEFENSIN_B.				
DE ANTIMICROBIAL PEPTIDE MGDPB.						GN				
OS Mytilis galloprovincialis (Mediterranean mussel).						OS Ornithodoros moubata (Soft tick).				
OC Eukaryota; Metazoa; Mollusca; Pteriomorphia; Mytiloids;						OC Eukaryota; Metazoa; Arthropoda; Cheliceraata; Arachnida; Acari;				
OC Mytiloidea; Mytilidae; Mytilus.						OC Parasitiformes; Ixodida; Argasidae; Ornithodoros.				
OX NCBI_TAXID=29158;						OX NCBI_TAXID=6338;				
RN [1]						RN [1]				
RP SEQUENCE FROM N.A.						RP SEQUENCE FROM N.A.				
RA Mitta G., Hubert F., Dyrinda E.A., Boudry P.; Roch P.;						RA Nakajima Y.;				
RT "Mytilin B and MG22, two antimicrobial peptides of marine mussels: gene structure and expression analysis.", Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.						RT "cDNA CLONING OF TICK DEFENSIN.";				
RT DR AB041815; BAB41027.1;						RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.				
DR InterPro; IPR001542; Artthro_defensin.						DR InterPro; IPR001542; Artthro_defensin.				
DR PROSITE; PS00425; ARTHROPOD_DEFENSINS; UNKNOWN_1.						DR PROSITE; PS00425; ARTHROPOD_DEFENSINS; UNKNOWN_1.				
DR SEQUENCE 73 AA; MW: 8290 MW; DEDBE90H06C9448D CRC64;						DR SEQUENCE 73 AA; MW: 8290 MW; DEDBE90H06C9448D CRC64;				
Query Match	44.8%;	Score 152;	DB 5;	Length 82;		Query Match	41.3%;	Score 140;	DB 5;	Length 73;
Best Local Similarity 48.3%;		Pred. No. 2.9e-12;				Best Local Similarity 40.3%;		Pred. No. 9e-11;		
Matches 28; Conservative 8; Mismatches 20;		Indels 2;	Gaps 2;			Matches 27; Conservative 11; Mismatches 19;		Indels 10;	Gaps 3;	
SEQUENCE 82 AA; 9307 MW; 6D8F0039EF675ECE CRC64;						QY 4 IAIIFIYLVAFICLE--DGI-----VEAGFGCPENAGKCHRCKSIRR-RGGFCRGTF 53				
DR PROSITE; PS00265; GLYCO_HORMONE_BETA_1; UNKNOWN_1;						Db 7 VAIIVVALVATMAQEVDVHDQEVSVPVRROYGCFFNQYQCHSHGIRGTSYKGGYCTGRF 66				
DR PROSITE; PS00425; ARTHROPOD_DEFENSINS; UNKNOWN_1;						QY 54 RFTCVCY 60				
DR PROSITE; PS00425; GLYCO_HORMONE_BETA_1; UNKNOWN_1;						Db 67 KOTCKY 73				
DR PROSITE; PS00425; ARTHROPOD_DEFENSINS; UNKNOWN_1;						RESULT 5				
DR PROSITE; PS00425; GLYCO_HORMONE_BETA_1; UNKNOWN_1;						Q9Y0B0	PRELIMINARY;	PRT;	57 AA.	
DR PROSITE; PS00425; GLYCO_HORMONE_BETA_1; UNKNOWN_1;						ID Q9Y0B0				
DR PROSITE; PS00425; GLYCO_HORMONE_BETA_1; UNKNOWN_1;						AC Q9Y0B0;				
DR PROSITE; PS00425; GLYCO_HORMONE_BETA_1; UNKNOWN_1;						DT 01-NOV-1999 (TREMBLrel. 12, Created)				
DR PROSITE; PS00425; GLYCO_HORMONE_BETA_1; UNKNOWN_1;						DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)				
DR PROSITE; PS00425; GLYCO_HORMONE_BETA_1; UNKNOWN_1;						DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)				
DR PROSITE; PS00425; GLYCO_HORMONE_BETA_1; UNKNOWN_1;						DE MGDL ANTIMICROBIAL PBPTIDE (FRAGMENT).				
DR PROSITE; PS00425; GLYCO_HORMONE_BETA_1; UNKNOWN_1;						OS Mytilus galloprovincialis (Mediterranean mussel).				
DR PROSITE; PS00425; GLYCO_HORMONE_BETA_1; UNKNOWN_1;						OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorphia; Mytiloida;				
DR PROSITE; PS00425; GLYCO_HORMONE_BETA_1; UNKNOWN_1;						OC Mytiloidea; Mytilidae; Mytilus.				
DR PROSITE; PS00425; GLYCO_HORMONE_BETA_1; UNKNOWN_1;						OX NCBI_TAXID=29158;				
DR PROSITE; PS00425; GLYCO_HORMONE_BETA_1; UNKNOWN_1;						RN [1]				
DR PROSITE; PS00425; GLYCO_HORMONE_BETA_1; UNKNOWN_1;						RP SEQUENCE FROM N.A.				
DR PROSITE; PS00425; GLYCO_HORMONE_BETA_1; UNKNOWN_1;						RA Mitta G., Hubert F., Roch P.;				
DR PROSITE; PS00425; GLYCO_HORMONE_BETA_1; UNKNOWN_1;						RT "Defensins: involvement in mussel defense.";				
DR PROSITE; PS00425; GLYCO_HORMONE_BETA_1; UNKNOWN_1;						RT Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.				
DR PROSITE; PS00425; GLYCO_HORMONE_BETA_1; UNKNOWN_1;						DR AF163337; AAD45117.1;				
DR PROSITE; PS00425; GLYCO_HORMONE_BETA_1; UNKNOWN_1;						DR InterPro; IPR001542; Artthro_defensin.				
DR PROSITE; PS00425; GLYCO_HORMONE_BETA_1; UNKNOWN_1;						DR Pfam; PF01097; Artthro_defensin.				
DR PROSITE; PS00425; GLYCO_HORMONE_BETA_1; UNKNOWN_1;						DR PROSITE; PS00425; ARTHROPOD_DEFENSINS; 1.				
DR PROSITE; PS00425; GLYCO_HORMONE_BETA_1; UNKNOWN_1;						FT NON_TER 1 1				
DR PROSITE; PS00425; GLYCO_HORMONE_BETA_1; UNKNOWN_1;						SQ SEQUENCE 57 AA; 6658 MW; 7B86B8623C979326 CRC64;				
Query Match	43.2%	Score 146.5;	DB 5;	Length 73;		Query Match	31.9%	Score 108;	DB 5;	Length 57;
Best Local Similarity 30%;						Best Local Similarity 30%;				
Matches 28; Conservative 8; Mismatches 20;						Matches 28; Conservative 8; Mismatches 20;				
SEQUENCE 82 AA; 9307 MW; 6D8F0039EF675ECE CRC64;						SEQUENCE 82 AA; 9307 MW; 6D8F0039EF675ECE CRC64;				



Best Local Similarity 28.8%; Pred. No. 0.0062; 21; Indels 29; Gaps 4;	DR InterPro; IPR002061; Scorpion_toxin.
Matches 23; Conservative 7; Mismatches 21; Indels 29; Gaps 4;	DR InterPro; IPR001219; Toxin.
Qy 6 IIFIVLVAFCT-----LEDGIVE-----AGFGCPENAKGCHRCKS 41	DR Pfam; PF00537; Toxin_3; 1.
Db 5 IVEFAFVAMCIAHSLAAPAPEALEASVTRQRKRLCDLSSFAEKGFAA - NSILCAHCLA 62	DR PRINTS; PR00284; TOXIN.
Qy 42 IRRGGFCRTERFTCYCR 61	DR PRODOM; PD000908; Scorpion_toxin; 1.
Db 63 IGRGGACQ -- NGVCVCR 79	DR SMART; SM00505; Knot1; 1.
SEQUENCE 85 AA; 9588 MW; 416CB3D72A8927ED CRC64;	SQ
Query Match 22.3%; Score 75.5; DB 5; Length 85;	Query Match 22.3%; Score 75.5; DB 5; Length 85;
Best Local Similarity 26.6%; Pred. No. 0.022;	Best Local Similarity 26.6%; Pred. No. 0.022;
Matches 17; Conservative 13; Mismatches 27; Indels 7; Gaps 3;	Matches 17; Conservative 13; Mismatches 27; Indels 7; Gaps 3;
QY 3 STALIFTWVAFCLIEDGIVEAGFGCPENAK--CHRHKSKIRRGFCR--GTERTC 57	QY 3 STALIFTWVAFCLIEDGIVEAGFGCPENAK--CHRHKSKIRRGFCR--GTERTC 57
Db 8 SLLALIITGVG -- SVRDGYIADDRCNPYFCGRNAYCDGECKKKNAESGICQWASKYGNAC 65	Db 8 SLLALIITGVG -- SVRDGYIADDRCNPYFCGRNAYCDGECKKKNAESGICQWASKYGNAC 65
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)	DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)	DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)	DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE DEFENSIN (FRAGMENT).	DE ALPHA_TOXIN PRECURSOR.
OS Aedes albopictus (Forest day mosquito).	OS Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;	OC Eubaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;	OC Buthoidea; Buthiidae; Buthus.
OC	OC
OX NCBI_TaxID=7160;	OX NCBI_TaxID=34649;
RN [1]	RN [1]
RP SEQUENCE FROM N.A.	RP SEQUENCE FROM N.A.
RA Melinda P.M., Kostas B., Scott O.L.;	RA Ye J.J., Chen J.J., Zuo X.J., Ji Y.
RT "Wolbachia neither induces nor suppresses transcripts encoding	RT Cloning and characterization of cDNA sequences encoding two novel
RT antimicrobial peptides."; DEFRINSIN.	RT alpha toxin precursors from the Chinese scorpion Buthus martensii.
RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.	RT Karsch;
RL EMBL; AY005473; AAF97983.1; -.	RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR HSSP; P10891; 1 ICA.	DR EMBL; AF288607; AAG0580.1; -.
DR InterPro; IPR00154; Arthro_defensin.	DR HSSP; P17728; IL01.
DR Pfam; PF003614; Knot1.	DR InterPro; IP003614; Knot1.
DR PR00271; DEFRINSIN.	DR InterPro; IPR002061; Scorpion_toxin.
DR SMART; SM00505; Knot1; 1.	DR PRINTS; PR00284; TOXIN.
DR PROSTIE; PS000425; ARTHROPOD_DEFENSINS; 1.	DR PRODOM; PD000908; Scorpion_toxin; 1.
FT NON_TER 1 57	DR SMART; SM00505; Knot1; 1.
FT NON_TER 57 57	DR SIGNAL 1 19
SQ SEQUENCE 57 AA; 6214 MW; A9F9DF5AE4CA86A7 CRC64;	FT CHAIN 20 85 POTENTIAL.
Query Match 22.7%; Score 77; DB 5; Length 57;	SQ SEQUENCE 85 AA; 9366 MW; 94ED463D6193FFEF CRC64;
Best Local Similarity 44.4%; Pred. No. 0.0095; 5; Mismatches 11; Indels 4; Gaps 2;	Query Match 22.3%; Score 75.5; DB 5; Length 85;
Matches 16; Conservative 5; Mismatches 11; Indels 4; Gaps 2;	Best Local Similarity 26.6%; Pred. No. 0.022;
QY 24 AGFGCPENAKGCHRCKSTRRGFCRGTFRTCYC 59	Matches 17; Conservative 12; Mismatches 28; Indels 7; Gaps 3;
Db 26 SGFGVGDSA -CAAHCIARRNGGYCNA - kIVCVC 57	QY 3 STALIFTWVAFCLIEDGIVEAGFGCPENAK--CHRHKSKIRRGFCR--GTERTC 57
SEQUENCE FROM N.A.	Db 8 SLLALIITGVG -- SVRDGYIADDRCNPYFCGRNAYCDGECKKKNAESGICQWASKYGNAC 65
RA Zhu S.;	QY 58 VCYR 61
RA Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.	Db 66 WCYR 69
RC TISSUE=VENOM GLAND;	RESULT 12
RC DR Q9BLM4	Q9BLM4
TISSUE=VENOM GLAND;	PRELIMINARY;
RC DR Q9BLM4	PRELIMINARY;
TISSUE=VENOM GLAND;	PRT; 85 AA.
RC DR 01-JUN-2001 (TREMBLrel. 17, Created)	DR 01-JUN-2001 (TREMBLrel. 17, Created)

	Matches	15;	Conservative	4;	Mismatches	13;	Indels	4;	Gaps	2;
DT	01-JUN-2001	(TREMBLrel.	17,	Last sequence update)						
DT	01-DEC-2001	(TREMBLrel.	19,	Last annotation update)						
DE	PUTATIVE TOXIN PRECURSOR.									
OS	Androctonus australis (Sahara scorpion).									
OC	Buthidae; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;									
OC	Buthidae; Buthidae; Androctonus.									
NCBI_TAXID	6858;									
[1]										
RN	SEQUENCE FROM N.A.									
RX	MEDLINE=21211344;	PubMed=11311249;								
RA	Ceard B., Martin-Eauclaire M.F., Bougis P.E.;									
RT	"Evidence for a position-specific deletion as an evolutionary link between long- and short-chain scorpion toxins.";									
RT	FEBS Lett. 494:246-248 (2001).									
RL	EMBL: AJ308440; CAC37321.1;									
DR	HSSP: P01484; IAH0.									
DR	InterPro: IPR001614; Knot1.									
DR	InterPro: IPR003061; Scorpion_toxin.									
DR	InterPro: IPR012119; Toxin.									
DR	PRINTS: PRO0284; TOXIN.									
DR	PRODOM: PD000908; Scorpion_toxin; 1.									
SMART:	SM00505; Knot1; 1.									
KW	SIGNAL.									
FT	1	19	POTENTIAL_TOXIN.							
FT	CHAIN	20	85	POTENTIAL_TOXIN.						
SQ	SEQUENCE	85 AA;	9523 MW;	4059A69D80E4F090 CRC64;						
Query Match	21.18;	Score 71.5;	DB 5;	Length 85;						
Best Local Similarity	27.3%	Score 71.5;	DB 5;	Length 85;						
Matches	18;	Conservative	12;	Mismatches	7;	Gaps	3;			
QY	1	MKSIAIIFIVLVAFCLIEDGIVAEFGCPFNAGK--CHRHCKSIRRGGFCR - -GTFR	T 55							
Db	6 MISALLFMTGV E--SRKGTYIVDKRNCTFFCGRNAYCNDCEKKGAESGYCQWASPYGN	63								
QY	56	TCVCYR	61							
.DB	64	ACCYCY	69							
RESULT	13									
061721	ID	061721	PRELIMINARY;		PRT;	102 AA.				
AC	061721;									
DT	01-AUG-1998	(TREMBLrel.	07,	Created)						
DT	01-AUG-1998	(TREMBLrel.	07,	Last sequence update)						
DT	01-DEC-2001	(TREMBLrel.	19,	Last annotation update)						
DE	DEFENSIN.									
GN										
OS	Anopheles gambiae (African malaria mosquito).									
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;									
OC	Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;									
OC	Anophelidae.									
OX	NCBI_TAXID=7165;									
RN	[1]									
RP	SEQUENCE FROM N.A.									
RC	STRAIN_KWA;									
RA	Eagleston P., Lu W., Zhao Y.; EMBL/GenBank/DBJ databases									
RA	EMBL: AF063402; AAC18575.1;									
DR	HSSP: P10891; IICCA2.									
DR	InterPro: IPR001542; Arthro_defensin.									
DR	InterPro: IPR003614; Knot1.									
DR	PRINTS: PRO0271; Arthro_defensin.									
DR	SMART: SM00505; Knot1; 1.									
DR	PROSITE: PS00425; ARTHROPOD_DEFENSTNS; 1.									
SQ	SEQUENCE	102 AA;	10681 MW;	628829B1620CDC05 CRC64;						
Query Match	20.68;	Score 70;	DB 5;	Length 102						
Best Local Similarity	41.78;	Score 70;	DB 5;	Length 102						
Matches	20;	Conservative	12;	Mismatches	8;	Gaps	3;			
QY	24	AGFGCBPFNAGKCHRICKSIRRGGFCR - -TFR	TVCV 59							
Db	69 SGFGVNNL-CAANCIARYRGYCNS-KAVCV 100									
RESULT	14									
Q9Y0B1	ID	Q9Y0B1	PRELIMINARY;		PRT;	103 AA.				
AC	Q9Y0B1;									
DT	01-NOV-1999 (TREMBLrel.	12,	Created)							
DT	01-NOV-1999 (TREMBLrel.	12,	Last sequence update)							
DR	MYTILIN B ANTIMICROBIAL PEPTIDE PRECURSOR.									
OS	Mytilus galloprovincialis (Mediterranean mussel).									
OC	Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorphia; Mytiloidae.									
OC	Mytilidae; Mytilidae; Mytilus.									
OX	NCBI_TAXID=29158;									
RN	[1]									
RP	SEQUENCE FROM N.A.									
RA	Dyrynda E.A., Boudry P., Roch P.;									
RT	Mytilus G., Hubert F., Dyrynda E.A., Boudry P., Roch P.;									
RT	Two antimicrobial peptides of marine mussels:									
RT	gene structure and expression analysis."									
RT	Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.									
RT	EMBL: AF162336; AAD45013.1; -.									
RT	EMBL: AF177540; AAD52661.1; -.									
SQ	SEQUENCE	103 AA;	11344 MW;	8879EB83098F2C891 CRC64;						
Query Match	20.68;	Score 70;	DB 5;	Length 103;						
Best Local Similarity	33.98;	Score 70;	DB 5;	Length 103;						
Matches	20;	Conservative	12;	Mismatches	8;	Gaps	3;			
QY	1	MKSIAIIFIVLVAFCLIEDGIVAEFGCPFNAGKCHRICKSIRRGGFCR - -TFR	TVCV 59							
Db	1 MKAAVILALAVAIL---AVHEEAASC--ASRKGHCRAR-RRCGYYVSVLRGRCYC 51									
RESULT	15									
Q9FR81	ID	Q9FR81	PRELIMINARY;		PRT;	72 AA.				
AC	Q9FR81;									
DT	01-MAR-2001 (TREMBLrel.	16,	Created)							
DT	01-MAR-2001 (TREMBLrel.	16,	Last sequence update)							
DE	DISEASE RESISTANCE PROTEIN 230.									
GN	DRR230.									
OS	Pisum sativum (garden pea).									
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;									
OC	Spermatophyta; Magnoliophyta; eudicots; core eudicots; Rosidae;									
OC	eurosid I; Fabales; Fabaceae; Papilionoideae; Vicieae; Pisum.									
OX	NCBI_TAXID=3888;									
RN	[1]									
RP	SEQUENCE FROM N.A.									
RC	STRAIN=NCV. GREENFEAST;									
RA	Savensstrand H., Brose M., Strid A.;									
RT	"Stress-induced disease resistance response protein 230 cDNA from Pisum sativum cv. Greenfeast."									
RT	Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.									
DR	EMBL: AF139018; AAC13285.1; -.									
DR	InterPro: IPR02118; Gamma-thionin.									
DR	InterPro: IPR003614; Knot1.									
DR	PFam: PF00304; Gamma-thionin; 1.									
DR	Prodrom: PD002594; Gamma-thionin; 1.									
DR	SMART: SM00505; Knot1; 1.									
SQ	SEQUENCE	72 AA;	8000 MW;	00FB605FED0920 CRC64;						

Query	Match	20.5%	Score	69.5	DB	10	Length	72
Best Local	Similarity	29.9%	Pred.	No.	0.11			
Matches	20;	Conservative	10;	Mismatches	24;	Indels	13;	Gaps
Qy	2	K\$TIA\$TIVLVAFACILED\$IVEAGFGCPFNAGK				-CHRICKSIRRRG-	--GFCR	50
Ddb	4	K\$IA\$CL\$FLLVLVIAQEVIVSEANTCENLAGSYKGVCFGGCDRHCRT-	-QEGATSGRCR	61				
Qy	51	GPFERTTC	57					
Ddb	62	DD\$FRWCW	68					

Search completed: September 17, 2002, 15:45:33  
Job time: 129 sec